SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

RATIONAL BOUNDS FOR RATIOS OF MODIFIED BESSEL FUNCTIONS*

INGEMAR NÅSELL†

Abstract. Double sequences of rational upper and lower bounds for the ratio $I_{\nu+1}(x)/I_\nu(x)$, $x>0$, $\nu>-\frac{1}{2}$ or $\nu>-1$, are established. The bounds are shown to converge, in certain cases monotonically, to the ratio $I_{\nu+1}(x)/I_\nu(x)$. A comparison with other approximations is made.

1. Introduction. The modified Bessel function of the first kind I_ν is considered on the domain $x>0$. It is real for ν real and it is positive if $\nu \geqq -1$. It was proved in 1965 by Soni [12] that

$$I_{\nu+1}(x)/I_\nu(x)<1, \qquad x>0, \quad \nu>-\tfrac{1}{2}.$$

The aim of this paper is to extend Soni's result. Specifically, we define in § 3 a double sequence of nonnegative rational functions $L_{\nu,k,m}(x)$ ($x>0$, $\nu>-1$, $k, m=0,1,2,\cdots$) and a double sequence of positive rational functions $U_{\nu,k,m}(x)$ ($x>0$, $k, m=0,1,2,\cdots$, $\nu>-\frac{1}{2}$ if $m=0$, $\nu>-1$ otherwise), and we prove that the functions $L_{\nu,k,m}(x)$ are lower bounds and the functions $U_{\nu,k,m}(x)$ are upper bounds of the ratio $I_{\nu+1}(x)/I_\nu(x)$.

These results and some properties of the bounds are derived in §§ 2 and 3 below. The ideas used in our derivation are indicated by the following remarks.

The relation

$$\frac{I_{\nu+1}(x)}{I_\nu(x)}=\left[\frac{2(\nu+1)}{x}+\frac{I_{\nu+2}(x)}{I_{\nu+1}(x)}\right]^{-1} \tag{1}$$

shows that if $F(\nu,x)$ is a positive upper (nonnegative lower) bound for $I_{\nu+1}(x)/I_\nu(x)$ ($x>0$, $\nu>-\frac{1}{2}$), then

$$\left[\frac{2(\nu+1)}{x}+F(\nu+1,x)\right]^{-1}$$

is a nonnegative lower (positive upper) bound for the same ratio ($x>0$, $\nu>-1$). By using this result one can generate the double sequence of bounds

$$U_{\nu,k,0}(x),\quad L_{\nu,k,0}(x),\quad U_{\nu,k,1}(x),\quad L_{\nu,k,1}(x),\cdots, \qquad k=0,1,2,\cdots,$$

from the sequence of upper bounds $U_{\nu,k,0}(x)$, $k=0,1,2,\cdots$.

The sequence of upper bounds $U_{\nu,k,0}(x)$ is found as follows. It is shown in § 2 that the function

$$g_\nu(x)=x^{-\nu}e^{-x}I_\nu(x),$$

is completely monotonic. By repeated application of the recurrence relations for $I_\nu(x)$ one can express the kth derivative of $g_\nu(x)$ in a form that contains modified

* Received by the editors July 8, 1975, and in revised form July 8, 1976.

† Department of Mathematics, Royal Institute of Technology, S-100 44 Stockholm 70, Sweden.
This work was supported by the Swedish Natural Science Research Council.

Bessel functions of the orders ν and $\nu+1$ only. The following explicit form is found from (3) and (7):

$$0<(-1)^k g_\nu^{(k)}(x)=2^k x^{-\nu} e^{-x} \frac{(\nu+1/2)_k}{(2\nu+1)_k} \cdot[\alpha_{\nu,k}(x) I_\nu(x)-\beta_{\nu,k}(x) I_{\nu+1}(x)],$$
$$x>0, \quad \nu>-\tfrac{1}{2}, \quad k=0,1,2,\cdots.$$

Here, $\alpha_{\nu,k}(x)$ and $\beta_{\nu,k}(x)$ are nonnegative polynomials in $1/x$ defined in (8) and (9), respectively. The sequence of upper bounds $U_{\nu,k,0}(x)$ of the ratio $I_{\nu+1}(x)/I_\nu(x)$ follows from the above expression; indeed

$$\frac{I_{\nu+1}(x)}{I_\nu(x)}<\frac{\alpha_{\nu,k}(x)}{\beta_{\nu,k}(x)}=U_{\nu,k,0}(x),$$
$$x>0, \quad \nu>-\tfrac{1}{2}, \quad k=0,1,2,\cdots.$$

Our work has been motivated by the need for bounds of modified Bessel functions in certain recent epidemiological models. The tropical parasitic infection schistosomiasis is transmitted by helminthic parasites. Males and females of the sexually mature forms of the parasite form pairs in blood vessels of human beings. Mathematical models have been formulated for the transmission of schistosomiasis in a community; see Nåsell [8] and Nåsell and Hirsch [9]. Modified Bessel functions appear in these models to account for monogamous mating between the parasites. Bounds for modified Bessel functions are needed in the analysis of the qualitative behaviors of the solutions of certain systems of nonlinear differential equations that appear in the models.

2. Some preparatory results. From Watson [13] we quote Schläfli's integral representation of Poisson's type for the modified Bessel function I_ν:

$$I_\nu(x)=\frac{(x/2)^\nu}{\Gamma(1/2)\Gamma(\nu+1/2)}\int_{-1}^{1}(1-t^2)^{\nu-1/2} e^{-xt}\,dt, \qquad \nu>-\tfrac{1}{2}.$$

By defining a function g_ν through the relation

$$g_\nu(x)=x^{-\nu} e^{-x} I_\nu(x), \qquad x>0, \quad \nu>-\tfrac{1}{2},$$

we find from the integral representation of I_ν that $g_\nu(x)$ is the Laplace transform of a function which is positive on the interval $0<t<2$ and equal to 0 for $t\geqq 2$. It follows that g_ν is completely monotonic on $(0,\infty)$ and that the strict inequalities

$$(-1)^k g_\nu^{(k)}(x)>0, \qquad x>0, \quad \nu>-\tfrac{1}{2}, \quad k=0,1,2,\cdots, \tag{2}$$

hold.

The recurrence relations for the modified Bessel function I_ν can be used to prove that

$$g_\nu^{(k)}(x)=(-2)^k x^{-\nu} e^{-x} \frac{(\nu+1/2)_k}{(2\nu+1)_k}[I_\nu(x)-G_{\nu,k}(x)], \qquad x>0, \quad \nu>-\tfrac{1}{2}, \quad k=0,1,2,\cdots, \tag{3}$$

where the function $G_{\nu,k}$ is defined as a linear combination of $I_{\nu+1}, \cdots, I_{\nu+k}$ by the relation

(4)
$$G_{\nu,k}(x)=\sum_{j=1}^{k}(-1)^{j+1}\binom{k}{j}\frac{2(2\nu+1)_{j-1}(\nu+j)}{(2\nu+k+1)_j}I_{\nu+j}(x),\qquad x>0,\quad \nu>-\tfrac{1}{2},\quad k=0,1,2,\cdots.$$

The function $G_{\nu,k}$ satisfies the recurrence relation

(5)
$$G_{\nu,k+1}(x)-G_{\nu,k}(x)=\frac{(2\nu+1)_2}{(2\nu+k+1)_2}[I_{\nu+1}(x)-G_{\nu+1,k}(x)],\qquad x>0,\quad \nu>-\tfrac{1}{2},\quad k=0,1,2,\cdots.$$

It follows from (2), (3) and (5) that

(6)
$$G_{\nu,k}(x)<G_{\nu,k+1}(x)<I_\nu(x),\qquad x>0,\quad \nu>-\tfrac{1}{2},\quad k=0,1,2,\cdots.$$

Thus the sequence of functions $\{G_{\nu,k}\}$ is a monotonically increasing sequence of lower bounds of the function I_ν. This result was proved in Nåsell [7] by using an expansion, given by Luke [6], of the confluent hypergeometric function in terms of modified Bessel functions.

We proceed to derive relations that express $I_{\nu+n}(x)$, where n is a nonnegative integer, in terms of $I_\nu(x)$ and $I_{\nu+1}(x)$. The relations take different forms, depending on whether n is even or odd. The relations contain certain polynomials in $1/x$ that are defined as follows:

$$a_{\nu,m}(x)=\sum_{k=0}^{m-1}\binom{m+k-1}{2k}(\nu+m-k+1)_{2k}\left(\frac{2}{x}\right)^{2k}+\delta(m),\qquad x\neq 0,\quad m=0,1,2,\cdots,$$

and

$$b_{\nu,m}(x)=\sum_{k=0}^{m-1}\binom{m+k}{2k+1}(\nu+m-k)_{2k+1}\left(\frac{2}{x}\right)^{2k+1},\qquad x\neq 0,\quad m=0,1,2,\cdots.$$

Here, $\delta(0)=1$ while $\delta(m)=0$ for $m\neq 0$.

The polynomials $a_{\nu,m}$ and $b_{\nu,m}$ are closely related. Specifically, the following relations hold:

$$a_{\nu,m+1}(x)-a_{\nu,m}(x)=(\nu+2m+1)(2/x)b_{\nu+1,m}(x),\qquad x\neq 0,\quad m=0,1,2,\cdots,$$

$$b_{\nu,m+1}(x)-b_{\nu,m}(x)=(\nu+2m+1)(2/x)a_{\nu-1,m+1}(x),\qquad x\neq 0,\quad m=0,1,2,\cdots.$$

Relations between I_ν, $I_{\nu+1}$, $I_{\nu+n}$ are given in Theorem 1.

THEOREM 1. *Let* $x\neq 0$ *and* $m=0,1,2,\cdots$. *Then*

(a) $I_{\nu+2m}(x)=a_{\nu,m}(x)I_\nu(x)-b_{\nu,m}(x)I_{\nu+1}(x)$,

(b) $I_{\nu+2m+1}(x)=a_{\nu-1,m+1}(x)I_{\nu+1}(x)-b_{\nu+1,m}(x)I_\nu(x)$.

The proof of this theorem follows by induction.

It is seen from (4) that the function $G_{\nu,k}$ is a linear combination of the functions $I_{\nu+1}, I_{\nu+2}, \cdots, I_{\nu+k}$. By applying Theorem 1 to (4) we are led to an expression for $I_\nu(x)-G_{\nu,k}(x)$ that involves polynomials in $1/x$ and $I_\nu(x)$, $I_{\nu+1}(x)$. The expression has the form

$$I_\nu(x)-G_{\nu,k}(x)=\alpha_{\nu,k}(x)I_\nu(x)-\beta_{\nu,k}(x)I_{\nu+1}(x), \qquad x>0, \quad \nu>-\tfrac{1}{2}, \quad k=0,1,2,\cdots. \tag{7}$$

Here, the polynomials $\alpha_{\nu,k}(x)$ and $\beta_{\nu,k}(x)$ are found from the relations

$$\begin{aligned}\alpha_{\nu,k}(x)=1+\sum_{m=1}^{[k/2]}\binom{k}{2m}\frac{2(2\nu+1)_{2m-1}(\nu+2m)}{(2\nu+k+1)_{2m}}a_{\nu,m}(x)\\ +\sum_{m=1}^{[(k-1)/2]}\binom{k}{2m+1}\frac{2(2\nu+1)_{2m}(\nu+2m+1)}{(2\nu+k+1)_{2m+1}}b_{\nu+1,m}(x),\\ x>0, \quad \nu>-\tfrac{1}{2}, \quad k=0,1,2,\cdots,\end{aligned} \tag{8}$$

and

$$\begin{aligned}\beta_{\nu,k}(x)=\sum_{m=0}^{[(k-1)/2]}\binom{k}{2m+1}\frac{2(2\nu+1)_{2m}(\nu+2m+1)}{(2\nu+k+1)_{2m+1}}a_{\nu-1,m+1}(x)\\ +\sum_{m=1}^{[k/2]}\binom{k}{2m}\frac{2(2\nu+1)_{2m-1}(\nu+2m)}{(2\nu+k+1)_{2m}}b_{\nu,m}(x),\\ x>0, \quad \nu>-\tfrac{1}{2}, \quad k=0,1,2,\cdots.\end{aligned} \tag{9}$$

3. Rational bounds for the ratio $I_{\nu+1}(x)/I_\nu(x)$. We proceed to apply the results of the preceding section in the derivation of rational upper and lower bounds of the ratio $I_{\nu+1}(x)/I_\nu(x)$. As a preparation we define two double sequences of polynomials in $1/x$ as follows:

$$A_{\nu,k,m}(x)=\alpha_{\nu+2m,k}(x)a_{\nu,m}(x)+\beta_{\nu+2m,k}(x)b_{\nu+1,m}(x), \qquad x>0, \quad k,m=0,1,2,\cdots, \quad \nu>-\tfrac{1}{2}-2m, \tag{10}$$

$$B_{\nu,k,m}(x)=\alpha_{\nu+2m,k}(x)b_{\nu,m}(x)+\beta_{\nu+2m,k}(x)a_{\nu-1,m+1}(x), \qquad x>0, \quad k,m=0,1,2,\cdots, \quad \nu>-\tfrac{1}{2}-2m. \tag{11}$$

Some elementary properties of these polynomials are summarized in Lemmas 1 and 2 below.

LEMMA 1. *Let* $x>0$, $k,m=0,1,2,\cdots$,

$$\nu>\begin{cases}-\tfrac{1}{2} & \text{if } m=0,\\ -2 & \text{otherwise.}\end{cases}$$

Then

(a) $A_{\nu,k,m}(x)\geqq 1$,

(b) $A_{\nu,k,m}(x)\to\infty$ *as* $k\to\infty$,

(c) $A_{\nu,k,m}(x)\to\infty$ *as* $m\to\infty$,

(d) $A_{\nu,k+1,m}(x)-A_{\nu,k,m}(x)>0$, $k+m>0$,

(e) $A_{\nu,k,m+1}(x)-A_{\nu,k,m}(x)>0,\ k+m>0,$

(f) $\begin{cases} A_{\nu,k,m}(x)=O(x^{-k-2m+2}) \text{ as } x\to 0,\ k+2m\geqq 2, \\ A_{\nu,1,0}(x)=O(x^{-1}) \text{ as } x\to 0, \end{cases}$

(g) $A_{\nu,k,m}(x)=O(1)$ *as* $x\to\infty$.

LEMMA 2. *Let* $x>0,\ k, m=0, 1, 2, \cdots,$

$$\nu>\begin{cases} -\frac{1}{2} & \text{if } m=0, \\ -1 & \text{otherwise.} \end{cases}$$

Then

(a) $B_{\nu,k,m}(x)>0,\ k+m>0,$

(b) $B_{\nu,k,m}(x)\to\infty$ *as* $k\to\infty$,

(c) $B_{\nu,k,m}(x)\to\infty$ *as* $m\to\infty$,

(d) $B_{\nu,k+1,m}(x)-B_{\nu,k,m}(x)>0,$

(e) $B_{\nu,k,m+1}(x)-B_{\nu,k,m}(x)>0,$

(f) $B_{\nu,k,m}(x)=O(x^{-k-2m+1})$ *as* $x\to 0,\ k+m>0,$

(g) $\begin{cases} B_{\nu,0,m}(x)=O(x^{-1}) \text{ as } x\to\infty,\ m>0, \\ B_{\nu,k,m}(x)=O(1) \text{ as } x\to\infty,\ k>0. \end{cases}$

We use the polynomials $A_{\nu,k,m}(x)$ and $B_{\nu,k,m}(x)$ to define two double sequences of rational functions as follows:

$$U_{\nu,k,m}(x)=\frac{A_{\nu,k,m}(x)}{B_{\nu,k,m}(x)}, \tag{12}$$

$$x>0,\quad k, m=0, 1, 2, \cdots,\quad \nu>\begin{cases} -\frac{1}{2} & \text{if } m=0, \\ -1 & \text{otherwise,} \end{cases}$$

$$L_{\nu,k,m}(x)=\frac{B_{\nu+1,k,m}(x)}{A_{\nu-1,k,m+1}(x)}, \tag{13}$$

$$x>0,\quad k, m=0, 1, 2, \cdots,\quad \nu>-1.$$

The explicit forms of these functions for $k+m\leqq 3$ are given in the Appendix. Note that $k=m=0$ gives $U_{\nu,0,0}(x)=\infty$ and $L_{\nu,0,0}(x)=0$, which are trivial upper and lower bounds, respectively, of $I_{\nu+1}(x)/I_\nu(x)$. With $k+m>0$ we prove in Theorems 2 and 3 below that each of the functions $U_{\nu,k,m}(x)$ is an upper bound of the ratio $I_{\nu+1}(x)/I_\nu(x)$, and that each of the functions $L_{\nu,k,m}(x)$ is a lower bound of the same ratio. The theorems also give monotonicity properties of the bounds, exhibit their asymptotic behaviors, and establish their convergence to the ratio $I_{\nu+1}(x)/I_{\nu(x)}$.

THEOREM 2. *Let* $x>0,\ k, m=0, 1, 2, \cdots,\ k+m>0,$

$$\nu>\begin{cases} -\frac{1}{2} & \text{if } m=0, \\ -1 & \text{otherwise.} \end{cases}$$

Then

(a) $\dfrac{I_{\nu+1}(x)}{I_\nu(x)} < U_{\nu,k,m}(x)$,

(b) $U_{\nu,k+1,m}(x) < U_{\nu,k,m}(x)$,

(c) $U_{\nu,0,m+1}(x) < U_{\nu,0,m}(x)$,

(d) $U_{\nu,k,m}(x) - \dfrac{I_{\nu+1}(x)}{I_\nu(x)} = O(x^{k+4m-1}), \qquad x \to 0,$

(e) $\begin{cases} U_{\nu,0,m}(x) - \dfrac{I_{\nu+1}(x)}{I_\nu(x)} = O(x), & x \to \infty, \\ U_{\nu,k,m}(x) - \dfrac{I_{\nu+1}(x)}{I_\nu(x)} = O(x^{-k}), & x \to \infty, \quad k > 0, \end{cases}$

(f) $\lim_{k\to\infty} U_{\nu,k,m}(x) = \dfrac{I_{\nu+1}(x)}{I_\nu(x)}$,

(g) $\lim_{m\to\infty} U_{\nu,k,m}(x) = \dfrac{I_{\nu+1}(x)}{I_\nu(x)}$.

Proof. Replacement of ν by $\nu+2m$ in relation (7) gives

$$I_{\nu+2m}(x) - G_{\nu+2m,k}(x) = \alpha_{\nu+2m,k}(x) I_{\nu+2m}(x) - \beta_{\nu+2m,k}(x) I_{\nu+2m+1}(x).$$

Here we apply relations (a) and (b) of Theorem 1 to express $I_{\nu+2m}(x)$ and $I_{\nu+2m+1}(x)$ in $I_\nu(x)$ and $I_{\nu+1}(x)$. By also introducing $A_{\nu,k,m}(x)$ and $B_{\nu,k,m}(x)$ from (10) and (11), respectively, we find that

$$(14) \qquad I_{\nu+2m}(x) - G_{\nu+2m,k}(x) = A_{\nu,k,m}(x) I_\nu(x) - B_{\nu,k,m}(x) I_{\nu+1}(x).$$

Now $I_\nu(x) > 0$ and $B_{\nu,k,m}(x) > 0$ by Lemma 2(a). Division of both sides of relation (14) by $B_{\nu,k,m}(x) I_\nu(x)$ gives, with the use of (12),

$$(15) \qquad U_{\nu,k,m}(x) - \frac{I_{\nu+1}(x)}{I_\nu(x)} = \frac{I_{\nu+2m}(x) - G_{\nu+2m,k}(x)}{B_{\nu,k,m}(x) I_\nu(x)}.$$

The denominator of the right hand side of (15) is positive by the argument above, and the numerator is positive by inequality (6). Thus $U_{\nu,k,m}(x)$ is an upper bound of $I_{\nu+1}(x)/I_\nu(x)$, and statement (a) holds.

The numerator of the right hand side of (15) decreases as k increases by 1, as seen from (6). The denominator of the right hand side of (15) increases as k increases by 1; see Lemma 2(d). Thus the upper bound $U_{\nu,k,m}(x)$ decreases as k increases by 1, as claimed in (b).

With $k=0$ we find from (4) that $G_{\nu,0}(x) = 0$ and that the numerator of the right hand side of (15) equals $I_{\nu+2m}(x)$. As m increases by 1 we find from the recurrence relation for I_ν that the numerator of the right hand side of (15) decreases. Lemma 2(e) shows that the denominator of the right hand side of (15) increases as m is increased by 1. Thus the upper bound $U_{\nu,0,m}(x)$ decreases as m is increased by 1, as claimed in (c).

The following asymptotic relations follow from results in [7]:

$$\frac{I_{\nu+n}(x)-G_{\nu+n,k}(x)}{I_\nu(x)}=O(x^n), \quad x\to 0, \quad n=0,1,2,\cdots,$$

and

$$\frac{I_{\nu+n}(x)-G_{\nu+n,k}(x)}{I_\nu(x)}=O(x^{-k}), \quad x\to\infty, \quad n=0,1,2,\cdots.$$

By using these results and the asymptotic behavior of $B_{\nu,k,m}(x)$ (Lemma 2(f), (g)), we find from (15) that (d) and (e) hold.

We note that $U_{\nu,1,0}(x)=1$. It follows therefore from (a) that $I_{\nu+n}(x)<I_\nu(x)$, $n=1,2,\cdots$. From this inequality and (6) we find that

$$(16) \qquad 0<\frac{I_{\nu+n}(x)-G_{\nu+n,k}(x)}{I_\nu(x)}\leqq 1, \qquad n=0,1,2,\cdots.$$

By applying this result and Lemma 2(b) to the right hand side of (15) we conclude that (f) holds. An application of (16) and Lemma 2(c) to the right hand side of (15) shows that (g) holds.

Some of the properties of the rational functions $L_{\nu,k,m}(x)$ are summarized in the following theorem.

THEOREM 3. *Let* $x>0$, $k, m=0,1,2,\cdots$, $k+m>0$, $\nu>-1$. *Then*

(a) $L_{\nu,k,m}(x)<\dfrac{I_{\nu+1}(x)}{I_\nu(x)}$,

(b) $L_{\nu,k,m}(x)<L_{\nu,k+1,m}(x)$,

(c) $L_{\nu,0,m}(x)<L_{\nu,0,m+1}(x)$,

(d) $\dfrac{I_{\nu+1}(x)}{I_\nu(x)}-L_{\nu,k,m}(x)=O(x^{k+4m+1})$, $\qquad x\to 0$,

(e) $\dfrac{I_{\nu+1}(x)}{I_\nu(x)}-L_{\nu,k,m}(x)=O(x^{-k})$, $\qquad x\to\infty$,

(f) $\lim_{k\to\infty} L_{\nu,k,m}(x)\searrow\dfrac{I_{\nu+1}(x)}{I_\nu(x)}$,

(g) $\lim_{m\to\infty} L_{\nu,k,m}(x)=\dfrac{I_{\nu+1}(x)}{I_\nu(x)}$.

This theorem follows in a manner similar to the proof of Theorem 2.

4. A comparison with other approximations. By writing each of the bounds $L_{\nu,k,m}(x)$ and $U_{\nu,k,m}(x)$ as a ratio of polynomials in x we find the degrees of the polynomials in numerator and denominator to be those given by Table 1.

TABLE 1

Function	Condition	Degree of numerator p	Degree of denominator q
$U_{\nu,0,m}$	$m \geqq 1$	$2m-1$	$2m-2$
$U_{\nu,k,m}$	$k \geqq 1, m \geqq 0$	$k+2m-1$	$k+2m-1$
$L_{\nu,0,m}$	$m \geqq 1$	$2m-1$	$2m$
$L_{\nu,k,m}$	$k \geqq 1, m \geqq 0$	$k+2m$	$k+2m$

Thus, our bounds are of the form

$$R_{p,q}(x)=\frac{S_p(x)}{T_q(x)},$$

where $S_p(x)$ is a polynomial in x of degree at most p and $T_q(x)$ is a polynomial in x of degree at most q. Following Baker [2], such a function is called a *Padé approximant of a function F* if its power series expansion agrees with that of $F(x)$ in its first $p+q+1$ terms, i.e. if

$$F(x)-R_{p,q}(x)=O(x^{p+q+1}), \qquad x\to 0.$$

From Table 1 and Theorems 2(d) and 3(d) we find that the bounds $L_{\nu,k,m}(x)$ and $U_{\nu,k,m}(x)$, $k+m>0$, are Padé approximants of $I_{\nu+1}(x)/I_\nu(x)$ for $k=0$ but not for $k>0$. The Padé approximants $L_{\nu,0,m}(x)$ and $U_{\nu,0,m}(x)$, $m>0$, are the approximants of the continued fraction expansion of $I_{\nu+1}(x)/I_\nu(x)$, which can be developed from relation (1). It has previously been observed by Amos [1] that rational bounds of $I_{\nu+1}(x)/I_\nu(x)$ can be derived from its continued fraction expansion.

Irrational bounds for $I_{\nu+1}(x)/I_\nu(x)$ have recently been given by Amos [1]. Putting

$$C_\nu(x)=\frac{x}{\nu+\frac{1}{2}+[x^2+(\nu+\frac{3}{2})^2]^{1/2}} \quad \text{and} \quad D_\nu(x)=\frac{x}{\nu+\frac{1}{2}+[x^2+(\nu+\frac{1}{2})^2]^{1/2}},$$

Amos shows that

$$0\leqq C_\nu(x)\leqq\frac{I_{\nu+1}(x)}{I_\nu(x)}\leqq D_\nu(x), \qquad x>0, \quad \nu\geqq 0.$$

We note that Amos' method can be used to give the more generally valid inequalities

$$0<C_\nu(x)<\frac{I_{\nu+1}(x)}{I_\nu(x)}, \qquad x>0, \quad \nu>-1$$

and

$$\frac{I_{\nu+1}(x)}{I_\nu(x)}<D_\nu(x), \qquad x>0, \quad \nu>-\tfrac{1}{2}.$$

Asymptotically we find that

$$D_\nu(x)-\frac{I_{\nu+1}(x)}{I_\nu(x)}=O(x^{-2}), \qquad x\to\infty,$$

$$D_\nu(x)-\frac{I_{\nu+1}(x)}{I_\nu(x)}=O(x), \qquad x\to 0,$$

$$\frac{I_{\nu+1}(x)}{I_\nu(x)}-C_\nu(x)=O(x^{-2}), \qquad x\to\infty,$$

$$\frac{I_{\nu+1}(x)}{I_\nu(x)}-C_\nu(x)=O(x^{3}), \qquad x\to 0.$$

A comparison with Theorem 2 shows that $D_\nu(x)$ is a sharper bound than $U_{\nu,1,0}(x)$ asymptotically both as $x\to 0$ and as $x\to\infty$, but that $U_{\nu,k,m}(x)$ is sharper than $D_\nu(x)$ asymptotically as $x\to 0$ and as $x\to\infty$ if $k\geqq 3$, $m\geqq 0$. In a similar manner we find from a comparison with Theorem 3 that $C_\nu(x)$ is a sharper bound than $L_{\nu,1,0}(x)$ asymptotically as $x\to 0$ and as $x\to\infty$, but that $L_{\nu,k,m}(x)$ is sharper than $C_\nu(x)$ asymptotically as $x\to 0$ and as $x\to\infty$ if $k\geqq 3$, $m\geqq 0$. The following lower bound for the ratio $I_\nu(x)/I_\mu(x)$ has been given by Ross in [11]:

$$\frac{I_\nu(x)}{I_\mu(x)}>\left(\frac{x}{2}\right)^{\nu-\mu}\frac{\Gamma(\mu+1/2)}{\Gamma(\nu+1/2)}, \qquad \mu>\nu>0, \quad x>0.$$

By taking $\mu=\nu+1$ in this inequality, one finds that

$$\frac{I_{\nu+1}(x)}{I_\nu(x)}<\frac{x}{2\nu+1}, \qquad x>0, \quad \nu>0.$$

Now

$$U_{\nu,0,1}(x)=\frac{x}{2(\nu+1)}.$$

It follows therefore from Theorem 2(c) that each of the inequalities

$$\frac{I_{\nu+1}(x)}{I_\nu(x)}<U_{\nu,0,m}(x), \qquad x>0, \quad \nu>-1, \quad m=1,2,\cdots,$$

is sharper and more general than that of Ross in the case $\mu=\nu+1$.

Soni's inequality is

$$\frac{I_{\nu+1}(x)}{I_\nu(x)}<U_{\nu,1,0}(x)=1, \qquad x>0, \quad \nu>-\tfrac{1}{2};$$

see [12].

It follows from Theorem 2(b) that each of the inequalities

$$\frac{I_{\nu+1}(x)}{I_\nu(x)}<U_{\nu,k,0}(x), \qquad x>0, \quad \nu>-\tfrac{1}{2}, \quad k=2,3,\cdots,$$

is sharper than Soni's inequality.

Other results that are stronger than Soni's inequality have been given by Jones [4], Lorch [5], Cochran [3], and Reudink [10]. Thus, Jones proved that if $\varepsilon > 0$, $\nu \geqq 0$, and $x > 0$, then the inequalities

$$0 < \frac{I_{\nu+\varepsilon}(x)}{I_\nu(x)} < 1$$

hold, and furthermore the ratio $I_{\nu+\varepsilon}(x)/I_\nu(x)$ has a positive derivative for $x > 0$, and it approaches the value 0 as $x \to 0$ and the value 1 as $x \to \infty$. Lorch used a comparison theorem to establish inequalities for Whittaker functions. As a special case of these results, he draws the same conclusions as Jones concerning the ratio $I_{\nu+\varepsilon}(x)/I_\nu(x)$ in the slightly more general situation when $\varepsilon > 0$, $\nu > -\frac{1}{2}\varepsilon$, $x > 0$. Cochran established the inequality

$$\frac{\partial I_\nu(x)}{\partial \nu} < 0, \qquad \nu \geqq 0, \quad x > 0,$$

and Reudink used a different method to prove Cochran's inequality for $\nu > 0$, $x > 0$.

Appendix. Explicit expressions for some of the rational bounds.

$$U_{\nu,0,0}(x) = \infty,$$

$$U_{\nu,1,0}(x) = 1,$$

$$U_{\nu,0,1}(x) = \frac{x}{2(\nu+1)},$$

$$U_{\nu,2,0}(x) = \frac{x}{\nu+(1/2)+x},$$

$$U_{\nu,1,1}(x) = \frac{2(\nu+2)x + x^2}{4(\nu+1)_2 + 2(\nu+1)x + x^2},$$

$$U_{\nu,0,2}(x) = \frac{4(\nu+2)_2 x + x^3}{8(\nu+1)_3 + 4(\nu+2)x^2},$$

$$U_{\nu,3,0}(x) = \frac{(\nu+1/2)x + 2x^2}{2(\nu+1/2)(\nu+1) + 3(\nu+1/2)x + 2x^2},$$

$$U_{\nu,2,1}(x) = \frac{2(\nu+2)(\nu+5/2)x + 2(\nu+2)x^2 + x^3}{4(\nu+1)_2(\nu+5/2) + 4(\nu+1)_2 x + 3(\nu+3/2)x^2 + x^3},$$

$$U_{\nu,1,2}(x) = \frac{8(\nu+2)_3 x + 4(\nu+2)_2 x^2 + 4(\nu+3)x^3 + x^4}{16(\nu+1)_4 + 8(\nu+1)_3 x + 12(\nu+2)_2 x^2 + 4(\nu+2)x^3 + x^4},$$

$$U_{\nu,0,3}(x) = \frac{16(\nu+2)_4 x + 12(\nu+3)_2 x^3 + x^5}{32(\nu+1)_5 + 32(\nu+2)_3 x^2 + 6(\nu+3)x^4},$$

$$L_{\nu,0,0}(x) = 0,$$

$$L_{\nu,1,0}(x) = \frac{x}{2(\nu+1)+x},$$

$$L_{\nu,0,1}(x)=\frac{2(\nu+2)x}{4(\nu+1)_2+x^2},$$

$$L_{\nu,2,0}(x)=\frac{(\nu+3/2)x+x^2}{2(\nu+1)(\nu+3/2)+2(\nu+1)x+x^2},$$

$$L_{\nu,1,1}(x)=\frac{4(\nu+2)_2x+2(\nu+2)x^2+x^3}{8(\nu+1)_3+4(\nu+1)_2x+4(\nu+2)x^2+x^3},$$

$$L_{\nu,0,2}(x)=\frac{8(\nu+2)_3x+4(\nu+3)x^3}{16(\nu+1)_4+12(\nu+2)_2x^2+x^4},$$

$$L_{\nu,3,0}(x)=\frac{2(\nu+3/2)(\nu+2)x+3(\nu+3/2)x^2+2x^3}{4(\nu+1)(\nu+3/2)(\nu+2)+6(\nu+1)(\nu+3/2)x+5(\nu+11/10)x^2+2x^3},$$

$$L_{\nu,2,1}(x)=\frac{4(\nu+2)_2(\nu+7/2)x+4(\nu+2)_2x^2+3(\nu+5/2)x^3+x^4}{8(\nu+1)_3(\nu+7/2)+8(\nu+1)_3x+8(\nu+2)(\nu+9/4)x^2+4(\nu+2)x^3+x^4},$$

$$L_{\nu,1,2}(x)=\frac{16(\nu+2)_4x+8(\nu+2)_3x^2+12(\nu+3)_2x^3+4(\nu+3)x^4+x^5}{32(\nu+1)_5+16(\nu+1)_4x+32(\nu+2)_3x^2+12(\nu+2)_2x^3+6(\nu+3)x^4+x^5},$$

$$L_{\nu,0,3}(x)=\frac{32(\nu+2)_5x+32(\nu+3)_3x^3+6(\nu+4)x^5}{64(\nu+1)_6+80(\nu+2)_4x^2+24(\nu+3)_2x^4+x^6}.$$

Acknowledgments. I thank Harold S. Shapiro for suggesting the comparison with Padé approximants and the referees for constructive comments on the presentation of these results.

REFERENCES

[1] D. E. Amos, *Computation of modified Bessel functions and their ratios*, Math. Comput., 28 (1974), pp. 239–251.

[2] G. A. Baker, *Essentials of Padé Approximants*, Academic Press, New York, 1974.

[3] J. A. Cochran, *The monotonicity of modified Bessel functions with respect to their order*, J. Math. and Phys., 46 (1967), pp. 220–222.

[4] A. L. Jones, *An extension of an inequality involving modified Bessel functions*, J. Math. and Phys., 47 (1968), pp. 220–221.

[5] L. Lorch, *Inequalities for some Whittaker functions*, Arch. Math. (Brno), 3 (1967), pp. 1–10.

[6] Y. L. Luke, *The Special Functions and their Approximations*, vol. II, Academic Press, New York, 1969.

[7] I. Nåsell, *Inequalities for modified Bessel functions*, Math. Comput., 28 (1974), pp. 253–256.

[8] ———, *Schistosomiasis in a community with external infection*, Proc. 8th Internat. Biometric Conf., Editura Academiei Republicii Socialiste Romania, 1975, pp. 123–131.

[9] I. Nåsell and W. M. Hirsch, *The transmission dynamics of schistosomiasis*, Comm. Pure Appl. Math., 26 (1973), pp. 395–453.

[10] D. O. Reudink, *On the signs of the ν-derivatives of the modified Bessel functions $I_\nu(x)$ and $K_\nu(x)$*, J. Res. Nat. Bur. Standards Sec. B, 72B (1968), pp. 279–280.

[11] D. K. Ross, *Inequalities for special functions*, SIAM Rev., 15 (1973), pp. 665–670.

[12] R. P. Soni, *On an inequality for modified Bessel functions*, J. Math. and Phys., 44 (1965), pp. 406–407.

[13] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, University Press, Cambridge, England, 1944.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

LIE THEORY AND THE WAVE EQUATION IN SPACE-TIME. 2. THE GROUP $SO(4, \mathbb{C})$*

E. G. KALNINS† AND WILLARD MILLER, JR‡

Abstract. Homogeneous solutions of the Laplace or wave equation in four complex variables correspond to eigenfunctions of the Laplace–Beltrami operator on the complex sphere $S_{3c}: \sum_{i=1}^4 z_i^2 = 1$. It is shown explicitly that variables separate in this eigenvalue equation for exactly 21 orthogonal coordinate systems, each system characterized by a pair of commuting symmetry operators in the enveloping algebra of $so(4, \mathbb{C})$. Standard group-theoretic methods are applied to derive generating functions and integral representations for the separated solutions. Henrici's theory of expansions in products of Legendre functions is incorporated into this more general scheme.

1. Introduction. In [1] we studied the relation between symmetry and separation of variables for the differential equation in 3 real variables satisfied by solutions of the wave equation $\partial_{tt}\Phi - \Delta_3\Phi = 0$ which are homogeneous of degree σ in x, y, z, t. The appropriate symmetry group was $SO(3, 1)$. Here we examine this relationship in the case where all variables are complex. Instead of the Hilbert space theory for expansions of solutions of the differential equation in terms of separable solutions as developed in [1] we here construct a theory of analytic expansions in terms of separable solutions.

We begin with the complex Laplace equation

$$\Delta_4\Phi(y) = 0, \qquad \Delta_4 = \sum_{j=1}^4 \partial_{y_j y_j},$$
$$\mathbf{y} = (y_1, y_2, y_3, y_4), \qquad y_j \in \mathbb{C}. \tag{1.1}$$

Clearly (1.1) is equivalent to the complex wave equation, (set $y_1 = x$, $y_2 = y$, $y_3 = z$, $y_4 = it$). We are interested in the solutions of (1.1) which are homogeneous of fixed degree $\sigma \in \mathbb{C}$: $\Phi(r\mathbf{y}) = r^\sigma \Phi(\mathbf{y})$. Introducing coordinates r, z_j such that $y_j = rz_j$, $\sum_{j=1}^4 z_j^2 = 1$ we see that these homogeneous functions are uniquely determined by their values on the complex unit sphere $S_{3c}: z_1^2 + z_2^2 + z_3^2 + z_4^2 = 1$. Indeed $\Phi(\mathbf{y}) = r^\sigma \Phi(\mathbf{z})$. The group $SO(4, \mathbb{C}) \equiv SO(4)$ has a natural action on S_{3c} which is determined by the Lie derivatives

$$I_{jk} = z_j \partial_{z_k} - z_k \partial_{z_j}, \qquad 1 \leqq j, k \leqq 4, \quad j \neq k.$$

(Since this paper deals with *local* Lie theory we are concerned only with the behavior of analytic functions in small neighborhoods of a given point. Thus $f(r) = r^\sigma$ can be defined precisely in a neighborhood of $r_0 \neq 0$ by choosing any branch of the global analytic function, e.g., if $r_0 = R_0 e^{i\varphi_0}$, $R_0 > 0$, $-\pi < \varphi_0 < \pi$ we can define $f(r)$ for $r = R e^{i\varphi}$ in a small neighborhood of r_0 by $f(r) = \exp(\sigma \ln R) e^{i\sigma\varphi}$. The branch chosen makes no difference in the computations to follow. However, in § 4 it is necessary to be more careful about domains of definition in order to determine precisely the regions of validity of our identities.

* Received by the editors April 1976, and in revised form July 23, 1976.
† Mathematics Department, University of Waikato, Hamilton, New Zealand.
‡ School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455.

In that section we use the above definition of r^σ for $r=R>0$ and extend by analytic continuation.)

It is straightforward to show from (1.1) that the restriction ψ of the homogeneous function Φ to S_{3c} satisfies the eigenvalue equation for the Laplace operator on S_{3c}:

$$(1.2)\qquad (I_{12}^2+I_{13}^2+I_{14}^2+I_{23}^2+I_{24}^2+I_{34}^2)\psi(\mathbf{z})=-\sigma(\sigma+2)\psi(\mathbf{z}).$$

Moreover, the symmetry algebra of (1.2) is $so(4)$, the Lie algebra of $SO(4)$. In other publications we have developed a method which relates the symmetry group of a linear partial differential equation to the possible coordinate systems in which the equation admits solutions via separation of variables, e.g., [2], [3]. Here the method is applied to (1.2).

In § 2 we apply results of Eisenhart [4] to construct all complex orthogonal coordinate systems in which (1.2) admits separation. We show that there are exactly twenty-one such systems. In § 3 we show that each system is characterized by a pair of commuting second-order operators $\mathcal{L}_1$, $\mathcal{L}_2$ in the enveloping algebra of $so(4)$ in the sense that the corresponding separable solutions are common eigenfunctions of these operators with the separation constants as eigenvalues. We also discuss the relationship between the subalgebras $so(3)$, $so(3)\times so(3)$ and $\mathscr{E}(2)$ of $so(4)$ and some of the simpler coordinate systems.

In § 4 it is shown how the Lie algebraic characterization of the separable solutions of (1.2) can be used to derive generating functions and addition theorems for these special functions. Since the basic theory of such expansions has been discussed elsewhere, [5], [6], we merely present a few of the most interesting cases.

Among the results is a new group theoretic proof of the addition theorem for Gegenbauer polynomials $C_n^\lambda(x)$. The standard group-theoretic proofs of this result, [7, Chap. 11], use global representations of the family of groups $SO(m)$ and are valid only for half-integer values of λ. The proof given here is much simpler, uses local representations of $SO(4)$ and is valid for general complex λ. In [8], Henrici gave simple, elegant proofs of this addition theorem and many other generating functions for products of Gegenbauer functions by employing complex variable techniques on the partial differential equation (4.17) below, an equation which is distinct from (1.2). We will show, that (4.17) is actually equivalent to (1.2) under the action of the conformal symmetry group $SO(6)$ of (1.1) and point out the underlying group structure of Henrici's technique. A related proof of the addition theorem which implicitly employs separation of variables can be found in a recent note by Koornwinder [9].

Finally, in § 5 we show how to construct integral representations for each of the twenty-one classes of separated solutions of (1.2) by transferring the action of $SO(4)$ from S_{3c} to S_{2c}.

We are ultimately concerned with the classification of all separable and R-separable complex coordinate systems for (1.1) and the study of all special functions which arise from the equation via separation of variables. The determination of all homogeneous orthogonal separable systems given here is a first step toward realization of this program.

Note that by characterizing each separable system in terms of Lie algebra generators we have to a considerable extent reduced problems concerning the expansion of one set of separable solutions in other sets to a problem in the representation theory of the symmetry algebra. In [1] we studied unitary representations and obtained Hilbert space expansions whereas here we study local representations and obtain analytic expansions.

2. Separation of variables for the Laplace operator on S_{3c}. Here we consider the problem of separation of variables for the equation $\Delta\psi=\sigma(\sigma+2)\psi$ where Δ is the Laplace operator on the complex sphere S_{3c}. This is not equivalent to the corresponding problem on the real sphere S_3 studied by Olevskii [10] and Eisenhart [4] since we allow the coordinates to be complex quantities and ignore the ranges of variations of the coordinates. We do, however, restrict ourselves to orthogonal coordinate systems. The method we use for evaluating the systems is an adaption of that used by Eisenhart for a space of constant curvature. Here we look for all complex solutions for the metric coefficients rather than for all real solutions as did Eisenhart.

Let $\{x_1, x_2, x_3\}$ be a complex analytic coordinate system on S_{3c}. If the system is orthogonal then the metric takes the form

$$ds^2=H_1^2\,dx_1^2+H_2^2\,dx_2^2+H_3^2\,dx_3^2 \tag{2.1}$$

and the equation $\Delta\psi=\sigma(\sigma+2)\psi$ in these coordinates reads

$$\frac{1}{H_1H_2H_3}\left[\left(\partial_{x_1}\left(\frac{H_2H_3}{H_1}\partial_{x_1}\psi\right)+\partial_{x_2}\left(\frac{H_1H_3}{H_2}\partial_{x_2}\psi\right)+\partial_{x_3}\left(\frac{H_1H_2}{H_3}\partial_{x_3}\psi\right)\right]\right. \\ =-\sigma(\sigma+2)\psi. \tag{2.2}$$

Eisenhart has shown that if (2.2) separates in the variables $\{x_1, x_2, x_3\}$ then the metric coefficients must have one of the forms

1. $H_1=1,\quad H_2=\phi(x_1),\quad H_3=\theta(x_1),$
2. $H_1=1,\quad H_2=\phi(x_1),\quad H_3=\phi(x_1)\theta(x_2),$
3. $H_1=1,\quad H_2^2=(x_2-x_3)X_2(x_2)\sigma_1^2(x_1),\quad H_3^2=(x_2-x_3)X_3(x_3)\sigma_1^2(x_1),$
4. $H_1^2=\sigma_1(x_1)+e\sigma_3(x_3),\quad H_2^2=\sigma_1(x_1)\sigma_3(x_3),$
   $H_3^2=\sigma_1(x_1)+e\sigma_3(x_3),\quad e=\pm1$
5. $H_i^2=(x_i-x_j)(x_i-x_k)X_i(x_i),\quad i\neq j\neq k\neq i.$

In addition to having one of these forms the metric coefficients H_i^2 must satisfy the requirement that the space have constant unit curvature. This condition is

$$\begin{aligned}&\frac{1}{H_j^2}\left(2\frac{\partial^2}{\partial x_j^2}\log H_i^2+\frac{\partial}{\partial x_j}\log H_i^2\frac{\partial}{\partial x_j}\log\frac{H_i^2}{H_j^2}\right)\\ &\quad+\frac{1}{H_i^2}\left(2\frac{\partial^2}{\partial x_i^2}\log H_j^2+\frac{\partial}{\partial x_i}\log H_j^2\frac{\partial}{\partial x_i}\log\frac{H_j^2}{H_i^2}\right)\\ &\qquad\qquad+\frac{1}{H_k^2}\frac{\partial}{\partial x_k}\log H_i^2\frac{\partial}{\partial x_k}\log H_j^2=-4,\end{aligned} \tag{2.3}$$

where i, j, k are distinct. We now compute the differential forms associated with the four types of metric and subject to constraints (2.3).

1. For metrics of type 1 we find from (2.3) for $i=1, j=2$ and $i=1, j=3$ that ϕ and θ satisfy the equation $d^2\psi/dx_1^2+\psi=0$, and for $i=2, j=3$ in (2.3) we have the constraint $(d\phi/dx_1)(d\theta/dx_1)+\phi\theta=0$. There are two distinct solutions:

$$\text{(i)}\quad \phi=\sin x_1,\quad \theta=\cos x_1,$$

$$\text{(ii)}\quad \phi=e^{ix_1},\quad \theta=e^{ix_1}.$$

The corresponding metrics are

$$(1)\quad ds^2=dx_1^2+\sin^2 x_1\, dx_2^2+\cos^2 x_1\, dx_3^2,$$

$$(2)\quad ds^2=dx_1^2+e^{2ix_1}(dx_2^2+dx_3^2),$$

2. For metrics of type 2 we find from (2.3) with $i=1, j=2$ and $i=1, j=3$ that $\phi''+\phi=0$. For $i=2, j=3$ we find $\theta''+(\phi^2+\phi'^2)\theta=0$. The possible solutions to these equations are

$$\text{(i)}\quad \phi=\sin x_1,\quad \theta=\sin x_2,$$

$$\text{(ii)}\quad \phi=\sin x_1,\quad \theta=e^{ix_2},$$

$$\text{(iii)}\quad \phi=e^{ix_1},\quad \theta=x_2.$$

The corresponding metrics are

$$(3)\quad ds^2=dx_1^2+\sin^2 x_1(dx_2^2+\sin^2 x_2\, dx_3^2),$$

$$(4)\quad ds^2=dx_1^2+\sin^2 x_1(dx_2^2+e^{2ix_2}\, dx_3^2),$$

$$(5)\quad ds^2=dx_1^2+e^{2ix_1}(dx_2^2+x_2^2\, dx_3^2).$$

3. For metrics of type 3 we find from (2.3) with $i=1, j=2$ and $i=1, j=3$ that $\sigma_1''+\sigma_1=0$. If $\sigma_1=\sin x_1$ then $H_1^2=1$, $H_2^2=(x_2-x_3)X_2\sin^2 x_1$ and $H_3^2=(x_2-x_3)X_3\sin^2 x_1$. For $i=2, j=3$ in (2.3) we obtain

$$2\left(\frac{1}{X_2}+\frac{1}{X_3}\right)+(x_2-x_3)\left[\left(\frac{1}{X_3}\right)'-\left(\frac{1}{X_2}\right)'\right]-4(x_2-x_3)^3=0.$$

Differentiation of this equation twice with respect to x_2 implies $(1/X_2)'''=-24$ so

$$1/X_2=-4x_2^3+bx_2^2+cx_2+d=f(x_2).$$

Similarly $X_3=-1/f(x_3)$. There are only three distinct systems of this type:

$$(6)\quad ds^2=dx_1^2+\sin^2 x_1(\operatorname{sn}^2(x_2,k)-\operatorname{sn}^2(x_3,k))(dx_2^2-dx_3^2),$$

$$(7)\quad ds^2=dx_1^2+\left(\frac{1}{\operatorname{ch}^2 x_2}-\frac{1}{\operatorname{ch}^2 x_3}\right)\sin^2 x_1(dx_2^2-dx_3^2),$$

$$(8)\quad ds^2=dx_1^2+\left(\frac{1}{x_3^2}-\frac{1}{x_2^2}\right)\sin^2 x_1(dx_2^2-dx_3^2).$$

Here, sn (x, k) is a Jacobi elliptic function and we adopt the notation sh x, ch x, th x for hyperbolic functions.

In these equations we have introduced new variables $\tilde{x}_j = \tilde{x}_j(x_j)$, $j = 2, 3$. In general, we do not distinguish between coordinate systems $\{x_1, x_2, x_3\}$ and $\{\tilde{x}_1, \tilde{x}_2, \tilde{x}_3\}$ where $\tilde{x}_j = \tilde{x}_j(x_j)$, $j = 1, 2, 3$.

If $\sigma_1 = e^{ix_1}$ and $i = 2$, $j = 3$, then (2.3) reduces to

$$2\left(\frac{1}{X_2}+\frac{1}{X_3}\right)+(x_2-x_3)\left[\left(\frac{1}{X_3}\right)'-\left(\frac{1}{X_2}\right)'\right]=0.$$

Differentiating this equation twice with respect to x_2 we find $(1/X_2)'''=0$ or $1/X_2 = ax_2^2+bx_2+c = h(x_2)$. Similarly $1/X_3 = -h(x_3)$. There are four distinct systems of this type:

$$(9)\quad ds^2 = dx_1^2 + e^{2ix_1}(\text{ch}^2\, x_2 - \text{ch}^2\, x_3)(dx_2^2 - dx_3^2),$$

$$(10)\quad ds^2 = dx_1^2 + e^{2ix_1}(e^{2x_2} + e^{2x_3})(dx_2^2 - dx_3^2),$$

$$(11)\quad ds^2 = dx_1^2 + e^{2ix_1}(x_2^2 + x_3^2)(dx_2^2 + dx_3^2),$$

$$(12)\quad ds^2 = dx_1^2 + e^{2ix_1}(4x_2 - 4x_3)(dx_2^2 - dx_3^2).$$

4. For metrics of type 4, equation (2.3) with $i = 1$, $j = 2$ yields the constraint

$$2\left(\sigma_1''-\frac{\sigma_1'^2}{\sigma_1}\right)+\sigma_3\left(2\frac{\sigma_1''}{\sigma_1}-\frac{\sigma_1'^2}{\sigma_1^2}\right)+\frac{\sigma_3'^2}{\sigma_3}=-4(\sigma_1+\sigma_3)^2.$$

Differentiating with respect to x_3 we obtain

$$\frac{\sigma_1''}{\sigma_1}-\frac{\sigma_1'^2}{\sigma_1^2}+\left(\frac{\sigma_3'^2}{\sigma_3}\right)'\frac{1}{\sigma_3'}=-8(\sigma_1+\sigma_3).$$

We can separate variables in this equation according to the scheme

$$2\frac{\sigma_1''}{\sigma_1}-\frac{\sigma_1'^2}{\sigma_1^2}+8\sigma_1=4c,$$

$$\left(\frac{\sigma_3'^2}{\sigma_3}\right)'\frac{1}{\sigma_3'}+8\sigma_3=-4c$$

where c is a separation constant. First integrals of these equations are

$$\sigma_1'^2 = 4\sigma_1(f + c\sigma_1 - \sigma_1^2),$$

$$\sigma_3'^2 = 4\sigma_3(f - c\sigma_3 - \sigma_3^2),$$

f is a constant. Choosing new variables $\hat{x}_1 = \sigma_1$, $\hat{x}_3 = -\sigma_3$ we obtain the metric

$$ds^2 = \frac{1}{4}(\hat{x}_1 - \hat{x}_3)\left[\frac{d\hat{x}_1^2}{\hat{x}_1(a-\hat{x}_1)(b-\hat{x}_1)} - \frac{d\hat{x}_3^2}{\hat{x}_3(a-\hat{x}_3)(b-\hat{x}_3)}\right] + \hat{x}_1\hat{x}_3\, dx_2^2,$$

where $ab = -f$, $a + b = c$. There are four distinct cases:

If $a \neq b$, $|a|, |b| > 0$, the metric can be reduced to

$$(13)\quad ds^2 = -k^2(\text{sn}^2(x_1, k) - \text{sn}^2(x_3, k))(dx_1^2 - dx_3^2) + \frac{k^2}{k'^2}\text{cn}^2(x_1, k)\,\text{cn}^2(x_3, k))\, dx_2^2, \qquad k' = \sqrt{1-k^2},$$

If $a=b\neq 0$ we find

$$(14)\quad ds^2=(\mathrm{th}^2 x_1-\mathrm{th}^2 x_3)(dx_1^2-dx_3^2)+\mathrm{th}^2 x_1\,\mathrm{th}^2 x_3\,dx_2^2,$$

while if $a=0$, $b\neq 0$, we obtain

$$(15)\quad ds^2=\left(\frac{1}{\mathrm{ch}^2 x_1}-\frac{1}{\mathrm{ch}^2 x_3}\right)(dx_1^2-dx_3^2)+\frac{1}{\mathrm{ch}^2 x_1\,\mathrm{ch}^2 x_3}dx_2^2.$$

Finally, if $a=b=0$ the metric becomes

$$(16)\quad ds^2=\left(\frac{1}{x_1^2}+\frac{1}{x_3^2}\right)(dx_1^2+dx_3^2)+\frac{1}{x_1^2x_3^2}dx_2^2.$$

5. For metrics of type 5, equation (2.3) with $i=1$, $j=2$ becomes

$$\frac{1}{X_3}+\frac{1}{(x_1-x_2)^2}\Bigg\{(x_3-x_2)^2\left[(x_1-x_3)\left(\frac{1}{X_1}\right)'-\left(\frac{2(x_3-x_1)}{x_2-x_1}+1\right)\frac{1}{X_1}\right]$$
$$+(x_3-x_1)^2\Bigg[(x_2-x_3)\left(\frac{1}{X_2}\right)'$$
$$-\left(\frac{2(x_3-x_2)}{x_1-x_2}+1\right)\frac{1}{X_2}\Bigg]\Bigg\}+4(x_3-x_1)^2(x_3-x_2)^2=0.$$

Differentiating this equation twice with respect to x_2 we obtain a polynomial of order three in x_3. The coefficient $g(x_1,x_2)$ of x_3^3 must be identically zero. Thus

$$\frac{\partial^2 g}{\partial x_2^2}=\left(\frac{1}{X_2}\right)^{(4)}+96=0$$

and $1/X_2=-4x_4^2+ax_2^3+bx_2^2+cx_2+d=f(x_2)$. Similarly $1/X_1=f(x_1)$ and $1/X_3=f(x_3)$. Five coordinate systems of this type can be distinguished. In each case the metric assumes the form

$$ds^2=\frac{(x_1-x_2)(x_1-x_3)}{f(x_1)}dx_1^2+\frac{(x_2-x_3)(x_2-x_1)}{f(x_2)}dx_2^2+\frac{(x_3-x_1)(x_3-x_2)}{f(x_3)}dx_3^2$$

and the systems are distinguished by the multiplicities of the zeros of $f(x)$. The distinct possibilities are

$$(17)\quad f(x)=-4(x-a)(x-b)(x-1)x,\quad a\neq b,$$

$$(18)\quad f(x)=-4(x-2)(x-1)x^2,$$

$$(19)\quad f(x)=-4(x-1)^2x^2,$$

$$(20)\quad f(x)=-4(x-1)x^3,$$

$$(21)\quad f(x)=-4x^4.$$

This completes the list of orthogonal coordinate systems on the complex sphere S_{3c} which permit separation of variables for the equation $\Delta\psi=\sigma(\sigma+2)\psi$. There are exactly 21 such systems.

3. Lie algebra characteristics of the separable systems. The three-dimensional complex sphere S_{3c} consists of those points (z_1, z_2, z_3, z_4) in complex four-dimensional Euclidean space such that $z_1^2+z_2^2+z_3^2+z_4^2=1$. The connected Lie subgroup of the complex Euclidean group which leaves this manifold invariant is $SO(4, \mathbb{C})$, the complex rotation group. A basis for the six-dimensional Lie algebra $so(4, \mathbb{C})$ of $SO(4, \mathbb{C})$ is

$$I_{kl} = z_k \partial_l - z_l \partial_k, \quad k, l = 1, 2, 3, 4, \quad k \neq l, \quad I_{kl} = -I_{lk}. \tag{3.1}$$

These basis elements satisfy the commutation relations

$$[I_{kl}, I_{st}] = \delta_{ls} I_{kt} - \delta_{ks} I_{lt} - \delta_{lt} I_{ks} + \delta_{kt} I_{ls}. \tag{3.2}$$

Further, if we put

$$\begin{aligned} J_1 &= \tfrac{1}{2}(I_{23}-I_{14}), \quad J_2 = \tfrac{1}{2}(I_{13}+I_{24}), \quad J_3 = \tfrac{1}{2}(I_{12}-I_{34}), \\ L_1 &= \tfrac{1}{2}(I_{23}+I_{14}), \quad L_2 = \tfrac{1}{2}(I_{13}-I_{24}), \quad L_3 = \tfrac{1}{2}(I_{12}+I_{34}), \end{aligned} \tag{3.3}$$

it becomes evident that $so(4, \mathbb{C}) \cong so(3, \mathbb{C}) \oplus so(3, \mathbb{C})$. Indeed

$$[J_i, J_j] = \varepsilon_{ijk} J_k, \quad [L_i, L_j] = \varepsilon_{ijk} L_k, \quad [J_i, L_j] = 0. \tag{3.4}$$

It can be verified by tedious computations that each of the 21 separable coordinate systems constructed in § 2 is characterized by a pair of commuting symmetric second-order operations $\mathcal{L}_1, \mathcal{L}_2$ in the enveloping algebra of $so(4, \mathbb{C})$. That is, the separable solutions $\psi = \psi_1(x_1)\psi_2(x_2)\psi_3(x_3)$ corresponding to such a system are characterized by the equations

$$\Delta\psi = \sigma(\sigma+2)\psi, \qquad \mathcal{L}_1\psi = \lambda_1\psi, \qquad \mathcal{L}_2\psi = \lambda_2\psi. \tag{3.5}$$

The eigenvalues λ_1, λ_2 are the separation constants. Expressed in terms of the generators of $so(4, \mathbb{C})$ the Laplace operator is

$$-\Delta = I_{12}^2 + I_{13}^2 + I_{14}^2 + I_{23}^2 + I_{24}^2 + I_{34}^2; \tag{3.6}$$

i.e., Δ is the Casimir operator for $so(4, \mathbb{C})$.

We now present the explicit coordinates and the corresponding operators $\mathcal{L}_1$, $\mathcal{L}_2$ for each of the 21 separable coordinate systems on S_{3c}.

(1) $z_1 = \sin x_1 \cos x_2, \quad z_2 = \cos x_1 \cos x_3,$

$z_3 = \cos x_1 \sin x_3, \quad z_4 = \sin x_1 \sin x_2,$

$\mathcal{L}_1 = I_{23}^2, \quad \mathcal{L}_2 = I_{14}^2;$

(2) $z_1 = \frac{1}{2}[e^{-ix_1} + (1+x_2^2+x_3^2)\, e^{ix_1}], \quad z_2 = ix_2\, e^{ix_1},$

$z_3 = ix_3\, e^{ix_1}, \quad z_4 = \dfrac{i}{2}[e^{-ix_1} + (-1+x_2^2+x_3^2)\, e^{ix_1}],$

$\mathcal{L}_1 = (I_{42}+iI_{21})^2, \quad \mathcal{L}_2 = (I_{34}+iI_{13})^2;$

(3) $z_1 = \sin x_1 \cos x_2, \quad z_2 = \sin x_1 \sin x_2 \cos x_3,$

$z_3 = \sin x_1 \sin x_2 \sin x_3, \quad z_4 = \cos x_1,$

$\mathcal{L}_1 = I_{12}^2 + I_{13}^2 + I_{23}^2, \quad \mathcal{L}_2 = I_{23}^2;$

(4) $z_1=\frac{1}{2}\sin x_1[e^{-ix_2}+(1-x_3^2)\,e^{ix_2}],\quad z_2=x_3\,e^{ix_2}\sin x_1,$

$z_3=\frac{-i}{2}\sin x_1[e^{-ix_2}-(1+x_3^2)\,e^{ix_2}],\quad z_4=\cos x_1,$

$\mathscr{L}_1=I_{12}^2+I_{13}^2+I_{23}^2,\quad \mathscr{L}_2=(I_{32}+iI_{21})^2;$

(5) $z_1=\frac{1}{2}[e^{-ix_1}+(1+x_2^2)\,e^{ix_1}],\quad z_2=i\,e^{ix_1}x_2\cos x_3,$

$z_3=i\,e^{ix_1}x_2\sin x_3,\quad z_4=(i/2)[e^{-ix_1}-(1-x_2^2)\,e^{ix_1}],$

$\mathscr{L}_1=(I_{42}+iI_{21})^2+(I_{34}+iI_{13})^2,\quad \mathscr{L}_2=I_{23}^2;$

(6) $z_1=\frac{1}{k'}\sin x_1\,\mathrm{dn}\,(x_2,k)\,\mathrm{dn}\,(x_3,k)\quad z_2=\frac{ik}{k'}\sin x_1\,\mathrm{cn}\,(x_2,k)\,\mathrm{cn}\,(x_3,k),$

$z_3=k\,\sin x_1\,\mathrm{sn}\,(x_1,k)\,\mathrm{sn}\,(x_3,k),\quad z_4=\cos x_1,$

$\mathscr{L}_1=I_{12}^2+I_{13}^2+I_{23}^2,\quad \mathscr{L}_2=I_{23}^2+k^2I_{13}^2;$

(7) $z_1=\frac{1}{2}\sin x_1\left(\frac{\mathrm{ch}\,x_3}{\mathrm{ch}\,x_2}+\frac{\mathrm{ch}\,x_2}{\mathrm{ch}\,x_3}\right),\quad z_2=\sin x_1\,\mathrm{th}\,x_2\,\mathrm{th}\,x_3,$

$z_3=i\,\sin x_1\left[\frac{1}{\mathrm{ch}\,x_2\,\mathrm{ch}\,x_3}-\frac{1}{2}\left(\frac{\mathrm{ch}\,x_3}{\mathrm{ch}\,x_2}+\frac{\mathrm{ch}\,x_2}{\mathrm{ch}\,x_3}\right)\right],\quad z_4=\cos x_1,$

$\mathscr{L}_1=I_{12}^2+I_{13}^2+I_{23}^2,\quad \mathscr{L}_2=-I_{12}^2-I_{13}^2+I_{23}^2+i\{I_{31},I_{32}\};$

(8) $z_1=\frac{-i\,\sin x_1}{8x_2x_3}[(x_3^2-x_2^2)^2+4],\quad z_2=\frac{\sin x_1}{2x_2x_3}[x_3^2+x_2^2],$

$z_3=\frac{\sin x_1}{8x_2x_3}[-(x_3^2-x_2^2)^2+4],\quad z_4=\cos x_1-\{I_{12},I_{13}\}+i\{I_{12},I_{23}\},$

$\mathscr{L}_1=I_{12}^2+I_{13}^2+I_{23}^2,\quad \mathscr{L}_2=-\{I_{12},I_{13}\}+i\{I_{12},I_{23}\};$

(9) $z_1=\frac{1}{2}(e^{-ix_1}+[1+\mathrm{ch}^2\,x_2+\mathrm{sh}^2\,x_3]\,e^{ix_1}),\quad z_2=i\,\mathrm{ch}\,x_2\,\mathrm{ch}\,x_3\,e^{ix_1},$

$z_3=\mathrm{sh}\,x_2\,\mathrm{sh}\,x_3\,e^{ix_1},\quad z_4=\frac{i}{2}(e^{-ix_1}+[-1+\mathrm{ch}^2\,x_2+\mathrm{sh}^2\,x_3]\,e^{ix_1}),$

$\mathscr{L}_1=(I_{42}+iI_{21})^2+(I_{34}+iI_{13})^2,\quad \mathscr{L}_2=I_{23}^2+(I_{34}+iI_{13})^2;$

(10) $z_1=\frac{1}{2}(e^{-ix_1}+[1+e^{2x_2}-e^{2x_3}]\,e^{ix_1}),\quad z_2=\frac{i}{\sqrt{2}}(\mathrm{sh}\,(x_2-x_3)+e^{x_2+x_3})\,e^{ix_1},$

$z_3=\frac{1}{\sqrt{2}}(\mathrm{sh}\,(x_2-x_3)-e^{x_2+x_3})\,e^{ix_1},$

$z_4=\frac{i}{2}(e^{-ix_1}+[-1+e^{2x_2}-e^{2x_3}]\,e^{ix_1},$

$\mathscr{L}_1=(I_{42}+iI_{21})^2+(I_{34}+iI_{13})^2,\quad \mathscr{L}_2=I_{23}^2-(I_{42}+I_{31}+i(I_{12}+I_{34}))^2;$

(11) $z_1=\frac{1}{2}(e^{-ix_1}+[1+\frac{1}{4}(x_2^2+x_3^2)^2]\,e^{ix_1}, \quad z_2=(i/2)(x_2^2-x_3^2)\,e^{ix_1},$

$z_3=ix_2x_3\,e^{ix_1}, \quad z_4=(i/2)(e^{-ix_1}+[-1+\frac{1}{4}(x_2^2+x_3^2)^2]\,e^{ix_1}),$

$\mathcal{L}_1=(I_{42}+iI_{21})^2+(I_{34}+iI_{13})^2, \quad \mathcal{L}_2=\{I_{23}, I_{42}+iI_{21}\};$

(12) $z_1=\frac{1}{2}(e^{-ix_1}+[1+2(x_2-x_3)^2(x_2+x_3)]\,e^{ix_1},$

$z_2=i[\frac{1}{2}(x_2-x_3)^2+(x_2+x_3)]\,e^{ix_1}, \quad z_3=[\frac{1}{2}(x_2-x_3)^2-(x_2+x_3)]\,e^{ix_1},$

$z_4=\frac{1}{2}(e^{-ix_1}+[-1+2(x_2-x_3)^2(x_2+x_3)]\,e^{ix_1}),$

$\mathcal{L}_1=(I_{42}+iI_{21})^2+(I_{34}+iI_{13})^2,$

$\mathcal{L}_2=\{I_{23}, I_{42}+I_{31}+iI_{21}+iI_{34}\}-i(I_{42}-I_{31}+i(I_{21}-I_{34}))^2;$

(13) $z_1=k\,\mathrm{sn}\,(x_1,k)\,\mathrm{sn}\,(x_3,k), \quad z_2=-i\frac{k}{k'}\,\mathrm{cn}\,(x_1,k)\,\mathrm{cn}\,(x_3,k)\cos x_2,$

$z_3=-i\frac{k}{k'}\,\mathrm{cn}\,(x_1,k)\,\mathrm{cn}\,(x_3,k)\sin x_2, \quad z_4=\frac{1}{k'}\,\mathrm{dn}\,(x_1,k)\,\mathrm{dn}\,(x_3,k),$

$\mathcal{L}_1=I_{23}^2, \quad \mathcal{L}_2=I_{12}^2+I_{13}^2+k^2I_{14}^2;$

(14) $z_1=\frac{1}{2}\left(\frac{\mathrm{ch}\,x_1}{\mathrm{ch}\,x_3}+\frac{\mathrm{ch}\,x_3}{\mathrm{ch}\,x_1}\right), \quad z_2=\mathrm{th}\,x_1\,\mathrm{th}\,x_3\,\mathrm{ch}\,x_2,$

$z_3=-i\,\mathrm{th}\,x_1\,\mathrm{th}\,x_3\,\mathrm{sh}\,x_2, \quad z_4=\frac{-i}{\mathrm{ch}\,x_1\,\mathrm{ch}\,x_3}+\frac{i}{2}\left(\frac{\mathrm{ch}\,x_1}{\mathrm{ch}\,x_3}+\frac{\mathrm{ch}\,x_3}{\mathrm{ch}\,x_1}\right),$

$\mathcal{L}_1=I_{23}^2, \quad \mathcal{L}_2=I_{24}^2+I_{34}^2-I_{12}^2-I_{13}^2-I_{14}^2-i\{I_{12}, I_{42}\}-i\{I_{13}, I_{43}\};$

(15) $z_1=\frac{-1}{2}\left(\frac{\mathrm{ch}\,x_3}{\mathrm{ch}\,x_1}+\frac{\mathrm{ch}\,x_1}{\mathrm{ch}\,x_3}\right)-\frac{x_2^2}{2\,\mathrm{ch}\,x_1\,\mathrm{ch}\,x_3}, \quad z_2=\frac{ix_2}{\mathrm{ch}\,x_1\,\mathrm{ch}\,x_3},$

$z_3=\mathrm{th}\,x_1\,\mathrm{th}\,x_3, \quad z_4=i\left[\frac{2-x_2^2}{2\,\mathrm{ch}\,x_1\,\mathrm{ch}\,x_3}-\frac{1}{2}\left(\frac{\mathrm{ch}\,x_1}{\mathrm{ch}\,x_3}+\frac{\mathrm{ch}\,x_3}{\mathrm{ch}\,x_1}\right)\right],$

$\mathcal{L}_1=(I_{42}+iI_{21})^2, \quad \mathcal{L}_2=2I_{12}^2+I_{13}^2+I_{14}^2-I_{34}^2+i(\{I_{12}, I_{42}\}+\{I_{13}, I_{43}\});$

(16) $z_1=\left[\frac{(x_1^2+x_3^2)^2+4}{8x_1x_3}+\frac{x_2^2}{2x_1x_3}\right], \quad z_2=\frac{-ix_2}{x_1x_3},$

$z_3=\frac{-i}{2}\left(\frac{x_1}{x_3}-\frac{x_3}{x_1}\right), \quad z_4=\frac{i(x_1^2+x_3^2)^2-4i}{8x_1x_3}+\frac{ix_2^2}{2x_1x_3},$

$\mathcal{L}_1=(I_{42}+iI_{21})^2, \quad \mathcal{L}_2=\{I_{32}, I_{42}+iI_{21}\}-\{I_{14}, iI_{34}-I_{13}\},$

(17) $z_1^2=\frac{-x_1x_2x_3}{ab}, \quad z_2^2=\frac{(x_1-1)(x_2-1)(x_3-1)}{(a-1)(b-1)},$

$z_3^2=\frac{-(x_1-b)(x_2-b)(x_3-b)}{(a-b)(b-1)b}, \quad z_4^2=\frac{(x_1-a)(x_2-a)(x_3-a)}{(a-b)(a-1)a},$

$\mathcal{L}_1=abI_{12}^2+aI_{13}^2+bI_{14}^2,$

$\mathcal{L}_2=(a+b)I_{12}^2+(a+1)I_{13}^2+(b+1)I_{14}^2+aI_{32}^2+bI_{42}^2+I_{43}^2;$

(18) $(iz_1+z_2)^2=\dfrac{x_1x_2x_3}{a},$

$$z_1^2+z_2^2=\frac{1}{a^2}[(a+1)x_1x_2x_3-a(x_1x_2+x_1x_3+x_2x_3)],$$

$$z_3^2=\frac{-(x_1-1)(x_2-1)(x_3-1)}{a-1},\quad z_4^2=\frac{(x_1-a)(x_2-a)(x_3-a)}{a^2(a-1)},$$

$$\mathscr{L}_1=(I_{42}-iI_{14})^2-a(I_{32}+iI_{13})^2-aI_{12}^2,$$

$$\mathscr{L}_2=(a+1)I_{12}^2+I_{14}^2+I_{42}^2-a(I_{13}^2+I_{32}^2)+(I_{42}+iI_{14})^2+(I_{32}+iI_{13})^2;$$

(19) $(z_1+iz_2)^2=-(x_1-1)(x_2-1)(x_3-1),$

$$z_1^2+z_2^2=2x_1x_2x_3-(x_1x_3+x_2x_3+x_1x_2)+1,\quad (z_3+iz_4)^2=-x_1x_2x_3,$$

$$z_3^2+z_4^2=x_1x_3+x_2x_3+x_1x_4-2x_1x_2x_3,$$

$$\mathscr{L}_1=2(I_{31}+iI_{32})^2+\{I_{31}+iI_{32},I_{24}+iI_{41}\}+I_{12}^2,$$

$$\mathscr{L}_2=2(I_{31}+iI_{32})^2+\{I_{31}+iI_{32},I_{24}+I_{41}\}-I_{34}^2;$$

(20) $(z_2-iz_1)^2=+x_1x_2x_3,\quad -2z_3(z_2-iz_1)=x_1x_2+x_1x_3+x_2x_3-x_1x_2x_3,$

$$z_1^2+z_2^2+z_3^2=+x_1x_2x_3-x_1x_2-x_1x_3-x_2x_3+x_1+x_2+x_3,$$

$$z_4^2=-(x_1-1)(x_2-1)(x_3-1),$$

$$\mathscr{L}_1=(I_{41}+iI_{42})^2+\{I_{32}-iI_{13},I_{12}\},$$

$$\mathscr{L}_2=I_{41}^2+I_{42}^2-I_{34}^2-(I_{41}+iI_{42})^2+\{I_{41}+iI_{42},I_{34}\};$$

(21) $(z_1+iz_2)^2=2x_1x_2x_3,\quad (z_1+iz_2)(z_3+iz_4)=-(x_1x_2+x_2x_3+x_1x_3),$

$$-(z_1+iz_2)(z_3-iz_4)+\tfrac{1}{2}(z_3+iz_4)^2=x_1+x_2+x_3,$$

$$\mathscr{L}_1=\tfrac{1}{2}\{I_{21},I_{14}+I_{23}+i(I_{31}+I_{24})\}-\tfrac{1}{4}[I_{13}+I_{24}+i(I_{23}+I_{41})]^2,$$

$$\mathscr{L}_2=\tfrac{1}{2}\{I_{21}+I_{43},I_{32}+I_{14}+i(I_{13}+I_{24})\}$$
$$+\tfrac{1}{2}\{I_{14}+I_{23}+i(I_{31}+I_{24}),I_{43}\}+\tfrac{1}{2}(I_{42}+iI_{23})^2-\tfrac{1}{2}(I_{13}+iI_{14})^2.$$

Here, $\{A,B\}=AB+BA$.

To understand the significance of these systems it is useful to examine some of the subalgebras of $so(4,\mathbb{C})$. As shown in (3.3) and (3.4) this algebra can be decomposed into $so(3,\mathbb{C})\oplus so(3,\mathbb{C})$, and it is easy to see that system (1) corresponds to this decomposition. Another $so(3,\mathbb{C})$ subalgebra of $so(4,\mathbb{C})$ has basis $\{I_{12},I_{13},I_{23}\}$ with commutation relations

$$[I_{12},I_{13}]=-I_{23},[I_{12},I_{23}]=I_{13},$$
$$[I_{13},I_{23}]=-I_{12}$$

and Casimir operator

$$I_{12}^2+I_{13}^2+I_{23}^2.$$

It is easily seen that the systems (3), (4), (6), (7), and (8) correspond to this Lie algebra reduction $so(4, \mathbb{C}) \supset so(3, \mathbb{C})$ and to coordinates on the sphere $S_{2c}: z_1^2 + z_2^2 + z_3^2 = \text{const}$. Indeed as indicated in [11] there are exactly five such systems corresponding to the $so(3, \mathbb{C})$ subalgebra.

The operators

$$E_1 = I_{42} + iI_{21}, \quad E_2 = I_{43} + iI_{31}, \quad E_3 = I_{23} \tag{3.7}$$

with commutation relations

$$[E_1, E_2] = 0, \qquad [E_1, E_3] = E_2, \qquad [E_2, E_3] = -E_1 \tag{3.8}$$

form a basis for the Euclidean subalgebra $\mathscr{E}(2, \mathbb{C})$ with invariant operator

$$E_1^2 + E_2^2. \tag{3.9}$$

The systems (2), (5), (9), (10), (11) and (12) correspond to the reduction $so(4, \mathbb{C}) \supset \mathscr{E}(2, \mathbb{C})$. Indeed, as shown in [12, Chap. 1], the complex Helmholtz equation with symmetry algebra $\mathscr{E}(2, \mathbb{C})$ separates in exactly six coordinate systems. The remaining nine of our twenty-one systems are not obviously related to subalgebra reductions. (However, systems (13), (14) involve the diagonalization of I_{23} and systems (15), (16) involve the diagonalization of E_1.)

Our separable systems can be understood from another viewpoint. In [13] we presented a group-theoretic analysis of the six separable systems for the Laplace operator on the real sphere S_3: $y_1^2 + y_2^2 + y_3^2 + y_4^2 = 1$. Here the symmetry algebra is $so(4, R)$. It is evident that each such real system can be analytically continued to a separable system on S_{3c}. Indeed the complexifications of these six systems correspond to our five complex systems (1), (3), (6), (13) and (17). (Elliptic cylindrical coordinates of types I and II complexify to the same system (13).) In [1] we analyzed the thirty-four separable systems for the Laplace operator on the hyperboloid $y_1^2 - y_2^2 - y_3^2 - y_4^2 = 1$ (symmetry algebra $so(3, 1)$). Complexification of the thirty-four systems yields all complex systems classified here with the exception of the systems (10), (12), (16) and the nonsubgroup systems (19), (21). However, it is evident by inspection that these five remaining cases arise by complexification of separable coordinates for the Laplace operator on the real hyperboloid $y_1^2 - y_2^2 + y_3^2 - y_4^2 = 1$, symmetry algebra $so(2, 2) \cong sl(2, R) \oplus sl(2, R)$. Thus all our complex separable coordinates are complexifications of real separable coordinates on the sphere S_3 and the hyperboloids $y_1^2 - y_2^2 - y_3^2 - y_4^2 = 1$, $y_1^2 - y_2^2 + y_3^2 - y_4^2 = 1$. Similarly the separated solutions are analytic continuations of the separated solutions for the real forms.

To be more specific, note that the coordinates corresponding to the subalgebra reduction $so(4, \mathbb{C}) \supset so(3, \mathbb{C})$ all have the form

$$z_1 = w_1 \sin x_1, \qquad z_2 = w_2 \sin x_1, \qquad z_3 = w_3 \sin x_1, \qquad z_4 = \cos x_1, \tag{3.10}$$

where $w_1^2 + w_2^2 + w_3^2 = 1$ and $w_j = w_j(x_2, x_3)$. The separated solutions are of the form

$$f(x_1, x_2, x_3) = \sin^l x_1 \mathscr{C}_{\sigma-l}^{l+1}(\cos x_1) h(x_2, x_3) \tag{3.11}$$

where $\mathscr{C}_\alpha^\lambda(s)$ is a solution of the Gegenbauer equation

$$(1 - s^2)\mathscr{C}_\alpha^{\lambda\prime\prime} + (2\lambda - 3)s\mathscr{C}_\alpha^{\lambda\prime} + \alpha(\alpha + 2\lambda)\mathscr{C}_\alpha^\lambda = 0 \tag{3.12}$$

and

$$(I_{12}^2+I_{13}^2+I_{23}^2)h=-l(l+1)h. \tag{3.13}$$

Similarly the coordinates corresponding to the reduction $so(4, \mathbb{C}) \supset \mathscr{E}(2, \mathbb{C})$ all have the form

$$\begin{aligned} z_1 &= \tfrac{1}{2}(e^{-ix_1}+[1+w_2^2+w_3^2]\, e^{ix_1}), \qquad z_2 = iw_2\, e^{ix_1}, \\ z_3 &= iw_3\, e^{ix_1}, \qquad z_4 = \frac{i}{2}(e^{-ix_2}+[-1+w_2^2+w_3^2]\, e^{ix_1}) \end{aligned} \tag{3.14}$$

where $w_j = w_j(x_2, x_3)$, $j = 2, 3$. The separated solutions are of the form

$$f(x_1, x_2, x_3) = e^{-ix_1} Z_{\pm(\sigma+1)}(i\omega\, e^{-ix_1})h(w_2, w_3), \tag{3.15}$$

where the cylindrical function $Z_\nu(s)$ is a solution of Bessel's equation

$$s^2 Z_\nu'' + sZ_\nu' + (s^2-\nu^2)Z_\nu = 0$$

and h is a solution of the complex Helmholtz equation

$$(\partial_{w_2w_2}+\partial_{w_3w_3}+\omega^2)h(w_2, w_3) = 0. \tag{3.16}$$

It follows from the above remarks that, except for the rather intractable systems (19) and (21), the separated solutions for all coordinate systems can be easily obtained by analytic continuation of results found in [1], [12] and [13].

4. Generating functions for the separated solutions. Here we are concerned with the analytic expansion of a particular separated solution of (2.2) in terms of a set of separated solutions. For the most part we shall confine our attention to expansions in terms of separated solutions corresponding to systems (1) and (3).

For system (1) with

$$\tau = \sin x_1\, e^{ix_2}, \qquad \xi = \cos x_1\, e^{ix_3}, \qquad w = \cos 2x_1, \tag{4.1}$$

one can easily verify that the functions

$$F^{(1)}_{\mu,m}(\tau, \xi, w) = {}_2F_1\left(\begin{matrix} \dfrac{m+\mu-\sigma}{2}, \ \dfrac{m+\mu+\sigma}{2}+1 \\ 1+\mu \end{matrix} \,\middle|\, \frac{1-w}{2}\right)\tau^\mu \xi^m \tag{4.2}$$

are solutions of (2.2) with

$$I_{14}F = i\mu F, \qquad I_{23}F = \text{im}\, F.$$

(For $(\sigma-\mu-m)/2 = n = 0, 1, 2, \cdots$ this solution is proportional to $P_n^{(\mu,m)}(w)\tau^m\xi^m$ where

$$P_n^{(\alpha,\beta)}(x) = \binom{n+\alpha}{n}\, {}_2F_1\left(\begin{matrix} -n, \alpha+\beta+n+1 \\ \alpha+1 \end{matrix} \,\middle|\, \frac{1-x}{2}\right) \tag{4.3}$$

is a Jacobi polynomial.) An independent solution is

$$(4.4)\qquad G^{(1)}_{\mu,m}(\tau,\xi,w)={}_2F_1\left(\begin{matrix}\dfrac{m+\mu-\sigma}{2},\dfrac{m+\mu+\sigma}{2}+1\\ 1+m\end{matrix}\middle|\dfrac{1+w}{2}\right)\tau^{\mu}\xi^{m}$$

if $(m+\mu\pm\sigma)/2$, m, μ are all noninteger.

For system (3) with

$$(4.5)\qquad \eta=-e^{ix_3}\sin x_2,\qquad \rho=-\cos x_2,\qquad q=\cos x_1,$$

it follows that the functions

$$(4.6)\qquad F^{(3)}_{l,m}(\eta,\rho,q)=\eta^{m}(1-q^2)^{l/2}C^{m+1/2}_{l-m}(\rho)C^{l+1}_{\sigma-l}(q)$$

satisfy (2.2) and

$$I_{23}F=\mathrm{im}\,F,\qquad (I_{12}^2+I_{13}^2+I_{23}^2)F=-l(l+1)F.$$

Here

$$C^{\nu}_{\alpha}(z)=\frac{\Gamma(\alpha+2\nu)}{\Gamma(\alpha+1)\Gamma(2\nu)}\,{}_2F_1\left(\begin{matrix}\alpha+2\nu,-\alpha\\ \nu+1/2\end{matrix}\middle|\frac{1-z}{2}\right)$$

ia a Gegenbauer function, a polynomial if $\alpha=0,1,2,\cdots$. An independent set of solutions is

$$(4.7)\qquad G^{(3)}_{l,m}(\eta,\rho,q)=\eta^{m}(1-q^2)^{l/2}C^{m+1/2}_{l-m}(\rho)D^{l+1}_{\sigma-l}(q),$$

where

$$D^{\nu}_{\alpha}(z)=e^{i\pi\nu}\frac{\Gamma(\alpha+2\nu)}{\Gamma(\nu)\Gamma(\alpha+\nu+1)}(2z)^{-\alpha-2}{}_2F_1\left(\begin{matrix}\nu+\alpha/2,\nu+\alpha/2+1/2\\ \nu+\alpha+1\end{matrix}\middle|z^{-2}\right).$$

The functions $C^{\nu}_{\alpha}(z)$, $D^{\nu}_{\alpha}(z)$ are analytic in the complex plane cut from -1 to $-\infty$ and from $+1$ to $-\infty$, respectively, along the real axis.

Now suppose $H(\tau,\xi,w)$, variables (4.1), is a solution of (2.2) which can be expanded in a convergent Laurent series in τ, ξ and is analytic in $1+w$ in a neighborhood of $w=-1$. Then it follows by Wiesner's principle, [5], [14] that

$$H(\tau,\xi,w)=\sum_{\mu,m}C_{\mu,m}G^{(1)}_{\mu,m}(\tau,\xi,w).$$

This is a generating function for the $\{G^{(1)}_{\mu,m}\}$. We can evaluate the constants $C_{\mu,m}$ by choosing special values for the variables. Similarly, if H is analytic in $1-w$ in a neighborhood of $w=1$ we can expand in terms of the basis $\{F^{(1)}_{\mu,m}\}$. Also, by making use of known expansion theorems for Gegenbauer polynomials, e.g., [15, p. 238], we can expand solutions of (2.2) in series of functions $\{F^{(3)}_{l,m}\}$ or $\{G^{(3)}_{l,m}\}$ where $l=\sigma-n$, $n=0,1,2,\cdots$.

A convenient way of constructing these generating functions H is to choose them to be separated solutions of (2.2) corresponding to one of our twenty-one coordinate systems. In this manner one can derive a wide variety of generating functions. However, the generating functions will usually lead to double sums. Here we limit ourselves to single-sum generating functions for the bases (1) and

(3) by restricting the generating functions to be eigenfunctions of I_{23} with eigenvalue im. Thus the sum over m can be omitted.

For example, expressing the solution (4.6) in terms of the variables (4.1) in the case $\sigma - l = n = 0, 1, 2, \cdots, l - m = k = 0, 1, 2, \cdots, \sigma \in \mathbb{C}$ and expanding in terms of the basis $\{G^{(1)}_{\mu,m}\}$ we find

$$\begin{aligned}(4.8)\quad &\tau^n[16\tau^2+(2\tau^2+w-1)^2]^{k/2}C_k^{\sigma-k-n+1/2}\left(\frac{w-1-2\tau^2}{\sqrt{16\tau^2+(2\tau^2+w-1)^2}}\right)\\ &C_n^{\sigma-n+1}\left(\frac{2\tau^2+w-1}{4\tau i}\right)=\sum_{s=0}^{n+k} b_s\tau^{2s}{}_2F_1\left(\begin{matrix}s-n-k, \sigma+s-n-k+1\\ \sigma-n-k+1\end{matrix}\,\middle|\,\frac{1+w}{2}\right).\end{aligned}$$

Setting $w=-1$ in this expression we obtain the generating function

$$\tau^n(-2\tau^2-2)^k\frac{(2\sigma-2k-2n)_k}{k!}C_n^{\sigma-n+1}\left(\frac{\tau^2-1}{2\tau i}\right)=\sum_{s=0}^{n+k} b_s\tau^{2s}$$

for the coefficients b_s. Similar but more complicated expansions can be obtained for $\sigma - l, l - m$ noninteger. Conversely, for μ and $(\sigma - m - \mu)/2 = n$ nonnegative integers we can expand the basis functions (4.2) in terms of the basis (4.6) to obtain

$$\begin{aligned}(4.9)\quad &(q-\rho\sqrt{q^2-1})^\mu P_n^{(\mu,m)}(-1+2(1-q^2)(1-\rho^2))\\ &=\sum_{s=0}^{2n+\mu} a_s(q^2-1)^{s/2}C_s^{m+1/2}(\rho)C_{2n+\mu-s}^{s+m+1}(q).\end{aligned}$$

Replacing ρ by $\xi(q^2-1)^{-1/2}$ and letting $q\to 1$ we find in the limit

$$(1-\xi)^\mu P_n^{(\mu,m)}(-1+2\xi^2)=\sum_{s=0}^{2n+\mu} a_s\binom{s+m-\frac{1}{2}}{s}\binom{2n+2m+\mu+s+1}{2n+\mu-s}\xi^s,$$

a simple generating function for the coefficients a_s. More generally, expanding a function (4.4) in terms of the basis functions (4.6) we find

$$\begin{aligned}(4.10)\quad &{}_2F_1\left(\begin{matrix}\frac{m+\mu-\sigma}{2}, \frac{m+\mu+\sigma}{2}+1\\ 1+m\end{matrix}\,\middle|\,(\rho^2-1)(q^2-1)\right)(q-\rho\sqrt{q^2-1})^\mu\\ &=\sum_{s=0}^{\infty} a_s(q^2-1)^{s/2}C_{\sigma-m-s}^{m+s+1}(q)C_s^{m+1/2}(\rho)\end{aligned}$$

valid for all ρ, q such that $|\rho\pm\sqrt{\rho^2-1}|>|(q-1)/(q+1)|^{1/2}$ and q is not pure imaginary. To compute the coefficients we set $\rho=\xi(q^2-1)^{-1/2}$ in (4.10) and let $q\to 1$:

$$\begin{aligned}&(1-\xi)^\mu{}_2F_1\left(\begin{matrix}\frac{m+\mu-\sigma}{2}, \frac{m+\mu+\sigma}{2}+1\\ 1+m\end{matrix}\,\middle|\,\xi^2\right)\\ &=\sum_{s=0}^{\infty} a_s\binom{\sigma+m+s+1}{\sigma-m-s}\binom{m+s-\frac{1}{2}}{s}(2\xi)^s, \qquad |\xi|<1.\end{aligned}$$

Since $[I_{14}, I_{23}]=0$ it follows that the function $\exp(\alpha I_{14})F^{(3)}_{l,m}$ is an eigenfunction of I_{23} with eigenvalue im. Thus one can expand this function in terms of the $\{F^{(3)}\}$ basis with only a single sum. Consider the case $m \in \mathbb{C}$, $l-m=n$, $\sigma-l=k$, n, $k=0,1,2,\cdots$. A straightforward computation yields

$$\exp(\alpha I_{14})F^{(3)}_{l,m}(\eta,\rho,q)=\eta^m C^m_n\left(\left[\frac{(1-q^2)(\rho^2-1)+1-h^2(\alpha)}{1-h^2(\alpha)}\right]^{1/2}\right)C^{m+n+1}_k(h(\alpha)) \cdot (1-h^2(\alpha))^{n/2}(1-q^2)^{m/2}$$

where

$$h(\alpha)=q\sqrt{1-\alpha^2}-\rho\alpha\sqrt{1-q^2}.$$

Thus,

$$(4.11)\quad C^{m+1/2}_n\left(\left[\frac{(1-q^2)(\rho^2-1)+1-h^2(\alpha)}{1-h^2(\alpha)}\right]^{1/2}\right)C^{m+n+1}_k(h(\alpha))(1-h^2(\alpha))^{n/2} = \sum_{s=0}^{n+k} a_s(\alpha)(1-q^2)^{s/2}C^{m+1/2}_s(\rho)C^{m+s+1}_{n+k-s}(q).$$

To obtain a simpler expression for the coefficients $a_s(\alpha)$ we set $\rho=\xi(1-q^2)^{-1/2}$ in (4.11) and let $q\to 1$:

$$C^{m+1/2}_n\left(\frac{\alpha+\xi\sqrt{1-\alpha^2}}{\sqrt{\alpha^2(1-\xi^2)+2\alpha\xi\sqrt{1-\alpha^2}}}\right)C^{m+n+1}_k(\sqrt{1-\alpha^2}-\alpha\xi) \cdot(\alpha^2(1-\xi^2)+2\alpha\xi\sqrt{1-\alpha^2})^{n/2} = \sum_{s=0}^{n+k} a_s(\alpha)\binom{m-\frac{1}{2}}{s}\binom{2m+n+k+s+1}{n+k-s}(2\xi)^s.$$

These expressions become much more tractable in the special case $n=0$. For that case and $t=\sqrt{1-\alpha^2}$ we see that the left-hand side of (4.11) is symmetric in q and t. Thus

$$a_s(t)=b_s(1-t^2)^{s/2}\,C^{m+s+1}_{k-s}(t)$$

and it is easy to check that

$$(4.12)\quad C^{m+1}_k(qt+\rho\sqrt{(1-q^2)(1-t^2)})=\frac{\Gamma(2m+1)}{[\Gamma(m+1)]^2} \cdot\sum_{s=0}^{k}\frac{2^{2s}(k-s)![\Gamma(m+s+1)]^2}{\Gamma(k+2m+s+2)}(2m+2s+1)[(1-q^2)(1-t^2)]^{s/2} \cdot C^{m+s+1}_{k-s}(q)C^{m+s+1}_{k-s}(t)C^{m+1/2}_s(\rho),\qquad m\in\mathbb{C}\quad k=0,1,2,\cdots$$

This is the addition theorem for Gegenbauer polynomials, [7, p. 178]. For $\sigma-l$ an arbitrary complex number one can obtain similar expansions for the bases $\{F^{(3)}\}$, $\{G^{(3)}\}$, [8].

Note that from the group-theoretic point of view, our last computation amounts to the determination of the matrix elements of the operator $\exp(\alpha I_{14})$ with respect to the basis $\{F^{(3)}_{l,m}\}$. Similarly one can compute the matrix elements of group operators $\exp(\sum_{i<j}\alpha_{ij}I_{ij})$ with respect to the $\{F^{(1)}\}$ and $\{G^{(1)}\}$ bases. Since these results are essentially contained in [12] and [16], we shall not reproduce them here.

For system (5) with

$$\tau = e^{ix_1}, \qquad r = x_2, \qquad \theta = x_3, \tag{4.13}$$

the functions

$$F^{(5)}_{\omega,m}(\tau, r, \theta) = \tau^{-1} e^{im\theta} J_{\sigma-1}(i\omega\tau^{-1}) J_m(r\omega) \tag{4.14}$$

satisfy (2.2) and

$$I_{23}F = imF, \qquad \mathcal{L}_1 F = -\omega^2 F.$$

Expanding $F^{(5)}_{\omega,m}$ in terms of the basis $\{G^{(1)}_{\mu,m}\}$ we obtain the identity ($\tau = t^{-1}$, $\beta = i\omega$, $\nu = -\sigma - 1$):

$$\begin{aligned} &t^{-\nu-m} J_\nu(\beta t) J_m\left(\beta t \sqrt{\frac{1+w}{2}}\right)\left(\frac{1+w}{2}\right)^{-m/2} \\ &\qquad = \sum_{s=0}^{\infty} a_s t^{2s} {}_2F_1\left(\begin{matrix} -s, -\nu-s \\ m+1 \end{matrix} \,\middle|\, \frac{1+w}{2}\right). \end{aligned} \tag{4.15}$$

To evaluate the coefficients a_s it is enough to set $w = -1$:

$$t^{-\nu} J_\nu(\beta t) \frac{(\beta/2)^m}{\Gamma(m+1)} = \sum_{s=0}^{\infty} a_s t^{2s}.$$

We see that (4.15) is equivalent to the well-known power series expansion for a product of Bessel functions [7, p. 11].

Expanding $F^{(5)}_{\omega,m}$ in terms of the basis $\{F^{(3)}_{l,m}\}$ we find

$$\begin{aligned} &(q-\rho\sqrt{q^2-1})^{-1} J_\nu\left(\frac{\omega}{q-\rho\sqrt{q^2-1}}\right) J_m\left(\frac{\sqrt{(q^2-1)(\rho^2-1)}}{q-\rho\sqrt{q^2-1}}\right) \\ &\quad \cdot (\rho^2-1)^{-m/2}(q^2-1)^{-m/2} \\ &\qquad = \sum_{s=0}^{\infty} b_s (q^2-1)^{s/2} C^{m+s+1}_{-\nu-m-s-1}(q) C^{m+1/2}_s(\rho) \end{aligned} \tag{4.16}$$

convergent for the same values of ρ, q as (4.10). As usual, a simpler generating function for the b_s can be obtained by setting $\rho = i\xi(1-q^2)^{-1/2}$ and letting $q \to 1$. A more complicated identity results when one expands $\exp(\alpha I_{14})F^{(5)}_{\omega,m}$ in terms of $\{F^{(3)}\}$ basis functions.

The expansions in terms of the $\{F^{(3)}\}$ basis listed above and various generalizations of these expansions are all treated in a beautiful paper by Henrici [8]. He studied the equation

$$\left(\partial_{xx} - \frac{(2\sigma+1)}{x}\partial_x + \partial_{yy} + \frac{(2m+1)}{y}\partial_y\right)\Phi(x, y) = 0 \tag{4.17}$$

which can be obtained from the complex Laplace equation $\Delta_4\psi = 0$ by separating off two variables, and showed that this equation admits R-separable solutions

$$(4.18)\qquad (\xi-\eta)^{m-\sigma}(\xi^2-1)^{(\sigma-l)/2}C_{\sigma-l}^{l+1}\left(\frac{\xi}{\sqrt{\xi^2-1}}\right)C_{l-m}^{m+1/2}(\eta),$$

$$\xi=\frac{1+ww^*}{2\sqrt{ww^*}},\qquad \eta=\frac{w+w^*}{2\sqrt{ww^*}},\qquad w=\frac{x+iy-c}{x+iy+c},\qquad w^*=\frac{x-iy-c}{x-iy+c},\qquad c \text{ const.}$$

He then developed an ingenious theory of expansions of analytic solutions of (4.17) in terms of the basis (4.18). Furthermore he observed that (4.17) permits separable solutions in coordinate systems analogous to (1) and (5) as well as (3) and derived generating functions for Gegenbauer functions by expanding each of these separated solutions as series in the basis (4.18).

Note that equation (4.17) and equation (2.2) with $I_{23}^2\Psi=-m^2\Psi$ each arise from the complex Laplace equation by separating off two variables. Moreover, in the next paper in this series we shall show that these two reduced equations are equivalent under the action of the local symmetry group $O(6, \mathbb{C})$ of the Laplace equation. Thus, every separable system for (2.2) is mapped to an R-separable system for (4.17) and conversely.

It follows that Henrici's analysis of (4.17) carries over to

$$(4.19)\qquad \Delta\Psi=\sigma(\sigma+2)\Psi,\qquad I_{23}^2\Psi=-m^2\Psi.$$

The local symmetry group of (4.19) consists of the operators $\exp(\alpha I_{14})$, $\alpha\in\mathbb{C}$, i.e., these operators map solutions into solutions. Thus if Ψ is a known analytic solution of (4.19) we can discuss the expansion of $\exp(\alpha I_{14})\Psi$ in terms of the bases $\{F^{(1)}\}$ and $\{F^{(3)}\}$. In Henrici's work, which concerns only expansions in the $\{F^{(3)}\}$ basis, this freedom is expressed by choosing a family of coordinate systems parametrized by a complex variable c. Systems corresponding to distinct values of c are equivalent under an appropriate symmetry operator $\exp(\alpha I_{14})$.

By inspection we see that (4.19) separates in five coordinate systems: (1), (3), (5), (13), (14). In his work on (4.17) Henrici employs R-separation in systems (1), (3) and (5), but he fails to note the R-separation in analogies of (13) and (14). (System (13) yields products of associated Lamé functions and will not be treated here. See, however, [1].)

For system (14) the functions

$$\begin{aligned}(4.20)\qquad & F_{\alpha,m}^{(14)}(u_1,x_2,u_3)\\ &=e^{mx_2}(1-u_1)^{\alpha-m/2}(1-u_3)^{\alpha-m/2}(u_1u_3)^{m/2}\\ &\quad\cdot {}_2F_1\left(\begin{matrix}\alpha+\sigma/2+1,\alpha-\sigma/2\\ m+1\end{matrix}\,\Big|\,u_1\right){}_2F_1\left(\begin{matrix}\alpha+\sigma/2+1,\alpha-\sigma/2\\ m+1\end{matrix}\,\Big|\,u_3\right),\\ &\qquad u_1=\operatorname{th}^2 x_1,\qquad u_3=\operatorname{th}^2 x_3\end{aligned}$$

satisfy (2.2) and

$$I_{23}F=imF,\qquad \mathcal{L}_2F=4(\alpha-m/2)^2F.$$

Expanding $F^{(14)}_{\alpha,m}$ in terms of the basis $\{G^{(1)}_{\mu,m}\}$ for $\alpha-\sigma/2=-n$, $n=0,1,2,\cdots$, we find

$$
\begin{aligned}
&{}_2F_1\left(\begin{matrix}\sigma-n+1,-n\\ m+1\end{matrix}\,\middle|\, u_1\right){}_2F_1\left(\begin{matrix}\sigma-n+1,-n\\ m+1\end{matrix}\,\middle|\, u_3\right)\\
(4.21)\qquad &=\sum_{s=0}^{n} b_s\tau^{2s}{}_2F_1\left(\begin{matrix}s-n,\sigma+s-n+1\\ m+1\end{matrix}\,\middle|\,\frac{1+w}{2}\right),\\
&u_{\substack{1\\3}}=\frac{w+3}{4}-\frac{\tau^2}{2}\mp\frac{1}{2}\left[\left(\frac{w+3}{2}-\tau^2\right)^2-(2w+2)\right]^{1/2}.
\end{aligned}
$$

Setting $w=-1$ we find

$$
\begin{aligned}
{}_2F_1\left(\begin{matrix}\sigma-n+1,-n\\ m+1\end{matrix}\,\middle|\,1-\tau^2\right)&=\frac{\Gamma(m+1)\Gamma(m+2n-\sigma)}{\Gamma(m+n+1)\Gamma(m+n-\sigma)}{}_2F_1\left(\begin{matrix}\sigma-n+1,-n\\ \sigma-2n-m+1\end{matrix}\,\middle|\,\tau^2\right)\\
&=\sum_{s=0}^{n} b_s\tau^{2s}.
\end{aligned}
$$

Similar but more complicated expressions can be obtained for $n\neq 0,1,2,\cdots$.

Expanding $F^{(14)}_{\alpha,m}$ in terms of the basis $\{F^{(3)}_{l,m}\}$ for $\alpha-\sigma/2=-n$, $2\alpha-m=k$, k, $n=0,1,2,\cdots$, we obtain

$$
\begin{aligned}
&[q-\rho\sqrt{q^2-1}]^k{}_2F_1\left(\begin{matrix}k+m+n+1,-n\\ m+1\end{matrix}\,\middle|\,u_1\right){}_2F_1\left(\begin{matrix}k+m+n+1,-n\\ m+1\end{matrix}\,\middle|\,u_3\right)\\
&\qquad=\sum_{s=0}^{k+2n} a_s(q^2-1)^{s/2}C^{m+s+1}_{k+2n-s}(q)C^{m+1/2}_s(\rho),\\
(4.22)\qquad &u_{\substack{1\\3}}=(1-\rho^2+\rho^2q^2+\rho q\sqrt{q^2-1})\\
&\qquad\pm[(1-\rho^2+\rho^2q^2+\rho q\sqrt{q^2-1})^2-(q^2-1)(\rho^2-1)]^{1/2}.
\end{aligned}
$$

A simpler generating function for the coefficients a_s can be found by setting $\rho=\xi(q^2-1)^{-1/2}$ and letting $q\to 1$.

For our final example we consider system (16) with basis functions

$$
\begin{aligned}
&F^{(16)}_{\lambda,n}(x_1,x_2,x_3)\\
(4.23)\qquad &=\exp[i\lambda x_2+\sqrt{\lambda}(x_3^2-x_1^2)/2](x_1x_3)^{\sigma+2}L_n^{(\sigma+1)}(\sqrt{\lambda}x_1^2)L_n^{(\sigma+1)}(\sqrt{\lambda}x_3^2)
\end{aligned}
$$

where $L_\nu^{(\alpha)}(x)$ is a generalized Laguerre function, a polynomial if $n=0,1,2,\cdots$, [17, p. 268]. These functions satisfy the operator equations

$$(I_{42}+iI_{21})F=i\lambda F,\qquad \mathscr{L}_2F=-2\sqrt{\lambda}(2n+\sigma+2)F.$$

Note that the operator $K=I_{34}+iI_{13}=(x_1^2+x_3^2)^{-1}(x_3\partial_{x_3}-x_1\partial_{x_1})$ commutes with $I_{42}+iI_{21}$. Thus the function $\exp(\alpha K)F^{(16)}_{n,k}(x_1,x_2,x_3)$, $k=0,1,2,\cdots$, can be expanded in a series of functions (4.23) with λ fixed and n running over the nonnegative integers. The result is

$$L_k^{(\sigma+1)}\left(\frac{x_1^2}{2}-\frac{x_3^2}{2}-\alpha+\mathcal{R}\right)L_k^{(\sigma+1)}\left(\frac{x_3^2}{2}-\frac{x_1^2}{2}+\alpha+\mathcal{R}\right)$$

$$(4.24)\qquad =\sum_{s=0}^{k} a_s L_s^{(\sigma+1)}(x_1^2)L_s^{(\sigma+1)}(x_3^2)$$

$$2\mathcal{R}=[(x_1^2-x_3^2-2\alpha)^2+4x_1^2x_3^2]^{1/2}.$$

(We choose the square root so that $2\mathcal{R}=x_1^2+x_3^2$ when $\alpha=0$.) For evaluation of the coefficients a_s it is enough to set $x_3=0$:

$$\binom{k+\sigma+1}{k}L_k^{(\sigma+1)}(x_1^2-2\alpha)=\sum_{s=0}^{k} a_s\binom{s+\sigma+1}{s}L_s^{(\sigma+1)}(x_1^2),$$

$$a_s=\binom{k+\sigma+1}{k}\binom{s+\sigma+1}{s}^{-1}L_{k-s}^{(-1)}(-2\alpha).$$

5. Integral representations for separated solutions. In analogy with a construction in [1] we can represent solutions of (2.2) as analytic functions on the complex sphere S_{2c}. Indeed, let $f(\mathbf{w})$ be analytic on $S_{2c}: w_1^2+w_2^2+w_3^2=1$, $w_1=(1-w_2^2-w_3^2)^{1/2}$ and let $F(\mathbf{z})$ be a function on S_{3c} defined by

$$(5.1)\qquad F(\mathbf{z})=\mathcal{I}[f]=\iint_{\mathcal{D}}[w_1z_1+w_2z_2+w_3z_3+iz_4]^{\sigma}f(\mathbf{w})\frac{dw_2\,dw_3}{w_1}$$

were $\mathcal{D}$ is a complex two-dimensional Riemann surface over w_2-w_3 space. We assume that the integration surface $\mathcal{D}$ and the analytic function f are chosen such that $\mathcal{I}[f]$ converges absolutely and arbitrary differentiation with respect to $z_1,\cdots,z_4$ is permitted under the integral sign. It follows that $F(\mathbf{z})$ is a solution of (2.2). (In fact F is a solution of the Laplace equation $\Delta_4F=0$ which is homogeneous of degree σ in $\mathbf{z}$.) Integrating by parts, we find that the operators I_{jk}, (3.1), acting on the solution space of (2.2) correspond to the operators

$$(5.2)\qquad \begin{aligned} &I_{12}=w_1\partial_{w_2}-w_2\partial_{w_1},\qquad I_{13}=-w_3\partial_{w_1},\qquad I_{23}=-w_3\partial_{w_2},\\ &I_{41}=-i(\sigma+2)w_1+i(1-w_1^2)\partial_{w_1}-iw_1w_2\partial_{w_2},\\ &I_{42}=-i(\sigma+2)w_2-iw_2w_1\partial_{w_1}+i(1-w_1^2)\partial_{w_2},\\ &I_{43}=-i(\sigma+2)w_3-iw_3w_1\partial_{w_1}-iw_3w_2\partial_{w_2}\end{aligned}$$

acting on the analytic functions $f(\mathbf{w})$, provided $\mathcal{D}$ and f are chosen such that the boundary terms vanish:

$$I_{jk}F=\mathcal{I}(I_{jk}f).$$

The point of this construction is that we can use the operators (5.2) to compute an eigenfunction $f_{\lambda\mu}$:

$$\mathcal{L}_1f=\lambda f,\qquad \mathcal{L}_2f=\mu f,$$

where $\mathcal{L}_1$, $\mathcal{L}_2$ are the operators characterizing one of the separable systems

(1)–(21). It follows that the integral $F_{\lambda\mu}=\mathscr{I}(f_{\lambda\mu})$ is a solution of (2.2) which satisfies

$$\mathscr{L}_1F=\lambda F, \qquad \mathscr{L}_2F=\mu F,$$

where now $\mathscr{L}_1, \mathscr{L}_2$ are expressed in terms of the operators (3.1). Thus F must be a separable solution of (2.2) in the coordinates to which $\mathscr{L}_1$, $\mathscr{L}_2$ correspond. This fact enables us to evaluate the integral to within a few normalization constants which are determined by inspection. Thus, this procedure leads to integral representations for the separable solutions of (2.2).

We illustrate the method with a single example treated in detail. We adopt complex coordinates α, η on S_{2c} such that

$$(5.3)\qquad \begin{aligned}&(w_1, w_2, w_3)=(\cos\alpha, \sin\alpha\cos\eta, \sin\alpha\sin\eta),\\ &\frac{dw_2\,dw_3}{w_1}=\sin\alpha\, d\alpha\, d\eta.\end{aligned}$$

These coordinates will prove useful in the construction of integral representations for separable systems in which the operator $I_{23}=\partial_\eta$ is diagonalized. If $f(\alpha, \eta)$ satisfies $I_{23}f=imf$ then $f=h(\alpha)t^m$ where $t=e^{i\eta}$. We choose the integration surface in the form $\mathscr{D}=C_1\times C_2$ where C_1 is the interval $[0, \pi]$ in the α-plane and C_2 is a simple closed curve surrounding the origin in the t-plane. Performing the t-integration and making use of the standard generating function for Gegenbauer polynomials [17, p. 175], we find

$$(5.4)\qquad \begin{aligned}F_m(z)=\mathscr{I}[f]&=-i\int_0^\pi h(\alpha)\oint\left[iz_4+z_1\cos\alpha+\frac{z}{2}\sin\alpha\left(\frac{u}{t}+\frac{t}{u}\right)\right]^\sigma t^{m-1}\,dt\,d\alpha\\ &=2\pi\left(\frac{z}{2}\right)^\sigma u^m\int_0^\pi(\sin\alpha)^{\sigma+1}C_{\sigma-m}^{-\sigma}\left[\frac{-z_1\cos\alpha-iz_4}{z\sin\alpha}\right]h(\alpha)\,d\alpha\end{aligned}$$

is a solution of (2.2) such that $I_{23}F=imF$. Here

$$z_2=z\left(\frac{u+u^{-1}}{2}\right),\qquad z_3=z\left(\frac{u-u^{-1}}{2i}\right),\qquad z_1^2+z_4^2+z^2=1$$

and we assume that σ, m are complex numbers such that $\sigma-m=n=0, 1, 2, \cdots$.

The requirement that $f(\alpha, \eta)=h(\alpha)t^m$ satisfy the system (3) eigenvalue equations

$$I_{23}f=imf, \qquad (I_{12}^2+I_{13}^2+I_{23}^2)f=-l(l+1)f$$

leads to a family of solutions

$$(5.5)\qquad h(\alpha)=(\sin\alpha)^m C_{l-m}^{m+1/2}(\cos\alpha).$$

Substituting this expression into (5.4) under the assumptions

$$\sigma-l=k, \qquad l-m=n, \qquad \sigma\in\mathbb{C}, \qquad k, n=0, 1, 2, \cdots,$$
$$\operatorname{Re} m>0, \qquad \operatorname{Re}(m+\sigma)>0$$

and using the fact that variables must separate in the resulting integral if coordinates (3) are employed, we obtain the identity

$$A C_k^{\sigma-k+1}(\cos x_1) C_n^{\sigma-k-n+1/2}(\cos x_2)$$

$$(5.6)\quad = (\sin x_1)^k (\sin x_2)^{k+n} \cdot \int_0^{\pi} (\sin\alpha)^{2\sigma-k-n+1}\, C_{k+n}^{-\sigma}(i \cot x_1 \csc x_2 \csc\alpha + \cot x_2 \cot\alpha) C_n^{\sigma-k-n+1/2}(\cos\alpha)\, d\alpha,$$

where A is a constant to be determined. To evaluate A we first let $x_2 \to 0$ and obtain

$$A C_n^{\sigma-k-n+1/2}(1) C_k^{\sigma-k+1}(\cos x_1)$$

$$(5.7)\quad = \frac{(\sin x_1)^k \Gamma(k+n-\sigma)}{\Gamma(-\sigma)(n+k)!} 2^{n+k} \cdot \int_0^{\pi} (\sin\alpha)^{2\sigma-2k-2n+1} C_n^{\sigma-k-n+1/2}(\cos\alpha)(i \cot x_1 + \cos\alpha)^{k+n}\, d\alpha,$$

an identity which is apparently due to Durand [18]. Finally, letting $x_1 \to 0$ and using the orthogonality relations for Gegenbauer polynomials we obtain

$$A = (-1)^{k+n} (i)^k 2^{2\sigma-k+1} \frac{\Gamma(\sigma+1)\Gamma(\sigma-k)}{\Gamma(2\sigma-k+2)}.$$

By varying the eigenfunctions (5.5) and the integration surface $\mathcal{D}$ one can find a variety of such identities. In each case the integral must separate in coordinates (3) and this permits easy evaluation. Similar remarks hold for each of the twenty-one separable systems.

REFERENCES

[1] E. G. KALNINS AND W. MILLER, *Lie theory and the wave equation in space–time. 1. The group SO(3, 1)*, J. Mathematical Phys., 18 (1977), pp. 1–16.

[2] E. G. KALNINS, *On the separation of variables for the Laplace equation* $\Delta\Psi + K^2\Psi = 0$ *in two- and three-dimensional Minkowski space*, this Journal, 6 (1975), pp. 340–374.

[3] W. MILLER, JR., *Symmetry, separation of variables and special functions*, Theory and Applications of Special Functions, Academic Press, New York, 1975.

[4] L. P. EISENHART, *Separable systems of Stackel*, Ann. Math, 35 (1934), pp. 284–305.

[5] W. MILLER, JR., *Lie Theory and Special Functions*, Academic Press, New York, 1968.

[6] C. P. BOYER, E. G. KALNINS, AND W. MILLER, *Symmetry and separation of variables for the Helmholtz and Laplace equations*, Nagoya Math. J., 60 (1976), pp. 35–80.

[7] A. ERDÉLYI, ET. AL., *Higher Transcendental Functions*, vol. 2, McGraw-Hill, New York, 1951.

[8] P. HENRICI, *Addition theorems for general Legendre and Gegenbauer functions*, J. Rational Mech. Anal., 4 (1955), pp. 983–1018.

[9] T. KOORNWINDER, *Three notes on classical orthogonal polynomials*, preprint Mathematisch Centrum, Amsterdam, 1975.

[10] P. M. OLEVSKII, *The separation of variables in the equation* $\Delta u + \lambda u = 0$ *for spaces of constant curvature in two and three dimensions*, Mat. Sb., 27 (1956), pp. 379–426, (in Russian).

[11] E. G. KALNINS AND W. MILLER, *Lie theory and separation of variables 11. The EPD equation*, J. Mathematical Phys., 17 (1976), pp. 369–377.

[12] W. MILLER, JR., *Symmetry and Separation of Variables*, Addison-Wesley, Reading, MA., to appear 1977.

[13] E. G. KALNINS, W. MILLER AND P. WINTERNITZ, *The group $O(4)$, separation of variables and the hydrogen atom*, SIAM J. Appl. Math., 30 (1976), pp. 630–664.

[14] L. WEISNER, *Group-theoretic origin of certain generating functions*, Pacific J. Math, 5 (1955), pp. 1033–1039.

[15] F. W. SCHÄFKE, *Einführing in die Theorie der Speziellen Funktionen der Mathematischen Physik*, Springer-Verlag, Berlin, 1963.

[16] W. MILLER, JR., *Lie theory and the Lauricella functions F_D*, J. Mathematical Phys., 13 (1972), pp. 1393–1399.

[17] A. ERDELYI, ET AL., *Higher Transcendental Functions*, vol. 1, McGraw-Hill, New York, 1951.

[18] L. DURAND, *Nicholson-type integrals for products of Gegenbauer functions and related topics*, Theory and Applications of Special Functions, Academic Press, New York, 1975.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

GENERALIZATION OF CERTAIN SUMMATIONS DUE TO RAMANUJAN*

CHIH-BING LING†

Abstract. This paper presents a method of summation of two groups of 16 series containing exponential or hyperbolic functions, which are generalized from certain summations due to Ramanujan. The series are summed in closed form by introducing two special coefficients when the parameter involved in the series takes on the value 1, $\sqrt{3}$ or $1/\sqrt{3}$.

1. Introduction. In this paper, summations of the following two groups of 16 series containing exponential or hyperbolic functions are considered. In Group I,

$$(1)\qquad \begin{aligned} &\begin{matrix} I_1 \\ I_2 \end{matrix} = \sum_{n=1}^{\infty} \frac{n^{2s-1}}{e^{2n\pi c} \mp 1}, &&\begin{matrix} I_3 \\ I_4 \end{matrix} = \sum_{n=1}^{\infty} \frac{(2n-1)^{2s-1}}{e^{(2n-1)\pi c} \mp 1}, \\ &\begin{matrix} I_5 \\ I_6 \end{matrix} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} n^{2s-1}}{e^{2n\pi c} \mp 1}, &&\begin{matrix} I_7 \\ I_8 \end{matrix} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(2n-1)^{2s}}{e^{(2n-1)\pi c} \mp 1}, \end{aligned}$$

and in group II,

$$(2)\qquad \begin{aligned} &II_1 = \sum_{n=1}^{\infty} \frac{n^{2s-1}}{\sinh n\pi c}, &&II_2 = \sum_{n=1}^{\infty} \frac{n^{2s}}{\cosh n\pi c}, \\ &II_3 = \sum_{n=1}^{\infty} \frac{(2n-1)^{2s-1}}{\sinh (2n-1)\pi c/2}, &&II_4 = \sum_{n=1}^{\infty} \frac{(2n-1)^{2s}}{\cosh (2n-1)\pi c/2}, \\ &II_5 = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} n^{2s-1}}{\sinh n\pi c}, &&II_6 = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} n^{2s}}{\cosh n\pi c}, \\ &II_7 = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(2n-1)^{2s}}{\sinh (2n-1)\pi c/2}, &&II_8 = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(2n-1)^{2s-1}}{\cosh (2n-1)\pi c/2}, \end{aligned}$$

where $s \leqq 0$ or 1 according to the exponent of n or $2n-1$ being $2s$ or $2s-1$, and $c = 1, \sqrt{3}$ or $1/\sqrt{3}$.

The following summations appear in Ramanujan's *Notebooks* [1] and *Collected Papers* [2] without proof:

$$(3)\qquad \begin{aligned} &\text{(i)}\ \sum_{n=1}^{\infty} \frac{n}{e^{2n\pi}-1} = \frac{1}{24} - \frac{1}{8\pi}, &&\text{(ii)}\ \sum_{n=1}^{\infty} \frac{n^5}{e^{2n\pi}-1} = \frac{1}{502}, \\ &\text{(iii)}\ \sum_{n=1}^{\infty} \frac{n^9}{e^{2n\pi}-1} = \frac{1}{264}, &&\text{(iv)}\ \sum_{n=1}^{\infty} \frac{n^{13}}{e^{2n\pi}-1} = \frac{1}{24}, \\ &\text{(v)}\ \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(2n-1)^{4s-1}}{\cosh (2n-1)\pi/2} = 0, &&\text{(vi)}\ \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(2n-1)^{6s-1}}{\cosh (2n-1)\pi\sqrt{3}/2} = 0, \end{aligned}$$

$$\text{(vii)}\ \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(2n-1)^{6s-1}}{\cosh (2n-1)\pi/2\sqrt{3}} = 0,$$

* Received by the editors April 30, 1976, and in revised form July 27, 1976.

† Department of Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061.

where $s \geqq 1$ in the last three series. The proofs of (i) and (iv) were supplied later by Watson [3] and Sandham and Cooper [4] and that of (v) by Rao [2, p. 326] and Sandham [5]. An obvious generalization of the first four series is

$$\sum_{n=1}^{\infty} \frac{n^{4s-3}}{e^{2n\pi}-1}, \qquad s \geqq 1. \tag{4}$$

A generalization of the remaining three series with respect to s is

$$\mathrm{II}_8 = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(2n-1)^{2s-1}}{\cosh (2n-1)\pi c/2}, \qquad s \geqq 1, \tag{5}$$

where $c = 1, \sqrt{3}$ or $1/\sqrt{3}$. A similar generalization of (4) with respect to s and also to the three values of c is

$$\mathrm{I}_1 = \sum_{n=1}^{\infty} \frac{n^{2s-1}}{e^{2n\pi c}-1}, \qquad s \geqq 1. \tag{6}$$

It is seen that (6) is the first series in Group I and (5) the last series in Group II. Further generalization extends methodically to the other series in each group. There is ground to assert that the list of series in the two groups is conclusive. A discussion will be given later.

The summations of the following four less extensively generalized series were considered by Sandham [5], for $s \geqq 0$:

$$\text{(7)} \quad \begin{aligned} &\text{(i)} \ \sum_{n=1}^{\infty} \frac{n^{4s+1}}{e^{2n\pi}-1}, &&\text{(ii)} \ \sum_{n=1}^{\infty} \frac{(2n-1)^{4s+1}}{e^{(2n-1)\pi}+1}, \\ &\text{(iii)} \ \sum_{n=1}^{\infty} \frac{(-1)^{n-1}n^{4s+1}}{\sinh n\pi}, &&\text{(iv)} \ \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(2n-1)^{4s+3}}{\cosh (2n-1)\pi/2}, \end{aligned}$$

which cover only one half of the series I_1, I_4, II_5 and II_8, respectively, in the case $c = 1$. In particular, the summation of the first series was also considered by Hardy [6] in using a different method. The series so generalized are summed by Sandham in closed form in terms of numerical constants which may or may not involve π. However, when they are further generalized into the two groups as shown in (1) and (2), it appears that summations in closed form are possible only when two special coefficients σ_4 and σ_6 are introduced. They are defined by the following double series:

$$\sigma_4 = \sum_{n,m=-\infty}^{\infty}{}' \frac{1}{(m+ni)^4}, \qquad \sigma_6 = \sum_{n,m=-\infty}^{\infty}{}' \frac{1}{(m+ne^{\pi i/3})^6}, \tag{8}$$

where the prime on the summation signs denotes the omission of simultaneous zeros of m and n from the double summation. These two coefficients can be expressed in terms of gamma functions [7]. Besides they have been evaluated numerically by the author to 221*S*.

2. Expansion of series into double series. It is known that the following four hyperbolic functions can be expanded into partial fractions [8] in the form:

$$(9)\qquad \begin{aligned} \coth \pi x &= \frac{1}{\pi x}+\frac{2x}{\pi}\sum_{m=1}^{\infty}\frac{1}{m^2+x^2},\\ \frac{1}{\sinh \pi x} &= \frac{1}{\pi x}+\frac{2x}{\pi}\sum_{m=1}^{\infty}\frac{(-1)^m}{m^2+x^2},\\ \tanh \pi x &= \frac{8x}{\pi}\sum_{m=1}^{\infty}\frac{1}{(2m-1)^2+4x^2},\\ \frac{1}{\cosh \pi x} &= -\frac{4}{\pi}\sum_{m=1}^{\infty}\frac{(-1)^m(2m-1)}{(2m-1)^2+4x^2}. \end{aligned}$$

Alternately, the functions can be expanded into power series of $e^{-\pi x}$ for positive values of x. Consequently, by equating the two expansions and further decomposing the partial fractions, we have

$$(10)\qquad \begin{aligned} \sum_{p=1}^{\infty} e^{-2p\pi x} &= -\frac{1}{2}+\frac{1}{2\pi x}+\frac{i}{2\pi}\sum_{m=1}^{\infty}\left(\frac{1}{m+ix}-\frac{1}{m-ix}\right),\\ \sum_{p=1}^{\infty} e^{-(2p-1)\pi x} &= \frac{1}{2\pi x}+\frac{i}{2\pi}\sum_{m=1}^{\infty}(-1)^m\left(\frac{1}{m+ix}-\frac{1}{m-ix}\right),\\ \sum_{p=1}^{\infty}(-1)^p e^{-2p\pi x} &= -\frac{1}{2}+\frac{i}{\pi}\sum_{m=1}^{\infty}\left(\frac{1}{2m-1+2ix}-\frac{1}{2m-1-2ix}\right),\\ \sum_{p=1}^{\infty}(-1)^p e^{-(2p-1)\pi x} &= \frac{1}{\pi}\sum_{m=1}^{\infty}(-1)^m\left(\frac{1}{2m-1+2ix}+\frac{1}{2m-1-2ix}\right). \end{aligned}$$

Next, manipulate both sides of the first three equations thus obtained by the following four kinds of operations before summing up from $n=1$ to ∞: (i) Differentiate $2s-1$ times and put $x=nc$. (ii) Differentiate $2s-1$ times and put $x=(2n-1)c/2$. (iii) Differentiate $2s-1$ times, put $x=nc$ and multiply by $(-1)^n$. (iv) Differentiate $2s$ times, put $x=(2n-1)c/2$ and multiply by $(-1)^n$. The double series formed on the left of each equation can readily be summed into a single series. We thus find from the first equation,

$$(11)\qquad \begin{aligned} \mathrm{I}_1 &= \frac{(2s-1)!}{(2\pi)^{2s}}\left[\frac{S_{2s}}{c^{2s}}+(-1)^s\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\frac{1}{(m+nci)^{2s}}+\frac{1}{(m-nci)^{2s}}\right\}\right],\\ \mathrm{II}_1 &= \frac{2(2s-1)!}{\pi^{2s}}\left[\frac{U_{2s}}{c^{2s}}+(-1)^s\right.\\ &\qquad \left.\cdot\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\frac{1}{\{2m+(2n-1)ci\}^{2s}}+\frac{1}{\{2m-(2n-1)ci\}^{2s}}\right\}\right],\\ \mathrm{I}_2 &= \frac{(2s-1)!}{(2\pi)^{2s}}\left[\frac{S^*_{2s}}{c^{2s}}-(-1)^s\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\frac{(-1)^n}{(m+nci)^{2s}}+\frac{(-1)^n}{(m-nci)^{2s}}\right\}\right],\\ \mathrm{II}_2 &= \frac{2(2s)!}{\pi^{2s+1}}\left[\frac{U^*_{2s+1}}{c^{2s+1}}-i(-1)^s\right. \end{aligned}$$

$$\cdot \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \left\{ \frac{(-1)^n}{\{2m+(2n-1)ci\}^{2s+1}} - \frac{(-1)^n}{\{2m-(2n-1)ci\}^{2s+1}} \right\} \bigg],$$

and from the second equation,

(12)

$$\mathrm{I}_3 = \frac{(2s-1)!}{2\pi^{2s}} \left[\frac{S_{2s}}{c^{2s}} + (-1)^s \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \left\{ \frac{(-1)^m}{(m+nci)^{2s}} + \frac{(-1)^m}{(m-nci)^{2s}} \right\} \right],$$

$$\mathrm{II}_3 = \frac{2^{2s}(2s-1)!}{\pi^{2s}} \bigg[\frac{U_{2s}}{c^{2s}} + (-1)^s$$

$$\cdot \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \left\{ \frac{(-1)^m}{\{2m+(2n-1)ci\}^{2s}} + \frac{(-1)^m}{\{2m-(2n-1)ci\}^{2s}} \right\} \bigg],$$

$$\mathrm{I}_4 = \frac{(2s-1)!}{2\pi^{2s}} \left[\frac{S^*_{2s}}{c^{2s}} - (-1)^s \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \left\{ \frac{(-1)^{m+n}}{(m+nci)^{2s}} + \frac{(-1)^{m+n}}{(m-nci)^{2s}} \right\} \right],$$

$$\mathrm{II}_4 = \frac{2^{2s+1}(2s)!}{\pi^{2s+1}} \bigg[\frac{U^*_{2s+1}}{c^{2s+1}} - i(-1)^s$$

$$\cdot \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \left\{ \frac{(-1)^{m+n}}{\{2m+(2n-1)ci\}^{2s+1}} - \frac{(-1)^{m+n}}{\{2m-(2n-1)ci\}^{2s+1}} \right\} \bigg],$$

where

(13)

$$S_{2s} = \sum_{n=1}^{\infty} \frac{1}{n^{2s}} = \frac{(2\pi)^{2s}}{2(2s)!} B_{2s}, \qquad s \geqq 1,$$

$$U_{2s} = \sum_{n=1}^{\infty} \frac{1}{(2n-1)^{2s}} = \left(1 - \frac{1}{2^{2s}}\right) S_{2s}, \qquad s \geqq 1,$$

$$S^*_{2s} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^{2s}} = \left(1 - \frac{1}{2^{2s-1}}\right) S_{2s}, \qquad s \geqq 1,$$

$$U^*_{2s+1} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)^{2s+1}} = \frac{E_{2s}}{2(2s)!} \left(\frac{\pi}{2}\right)^{2s+1}, \qquad s \geqq 0.$$

B_{2s} and E_{2s} are Bernoulli and Euler numbers, respectively. The first few values of each are $B_2 = 1/6$, $B_4 = 1/30$, $B_6 = 1/42$ and $E_0 = 1$, $E_2 = 1$, $E_4 = 5$.

Similarly, we find from the third equation,

(14)

$$\mathrm{I}_5 = -\frac{(-1)^s(2s-1)!}{\pi^{2s}} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \left\{ \frac{1}{(2m-1+2nci)^{2s}} + \frac{1}{(2m-1-2nci)^{2s}} \right\},$$

$$\mathrm{II}_5 = -\frac{2(-1)^s(2s-1)!}{\pi^{2s}}$$

$$\cdot \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \left[\frac{1}{\{2m-1+(2n-1)ci\}^{2s}} + \frac{1}{\{2m-1-(2n-1)ci\}^{2s}} \right],$$

$$\mathrm{I}_6 = \frac{(-1)^s(2s-1)!}{\pi^{2s}} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \left\{ \frac{(-1)^n}{(2m-1+2nci)^{2s}} + \frac{(-1)^n}{(2m-1-2nci)^{2s}} \right\},$$

$$\mathrm{II}_6=\frac{2i(-1)^s(2s)!}{\pi^{2s+1}}$$

$$\cdot\sum_{n-1}^{\infty}\sum_{m=1}^{\infty}\left[\frac{(-1)^n}{\{2m-1+(2n-1)ci\}^{2s+1}}-\frac{(-1)^n}{\{2m-1-(2n-1)ci\}^{2s+1}}\right].$$

Lastly, manipulate both sides of the fourth equation by the four kinds of operations as before except that the differentiation now is $2s$ times in (i), (ii), and (iii) and $2s-1$ times in (iv). We find after summation,

(15)

$$\mathrm{I}_7=-\frac{(-1)^s(2s)!2^{2s}}{\pi^{2s+1}}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\frac{(-1)^m}{(2m-1+2nci)^{2s+1}}+\frac{(-1)^m}{(2m-1-2nci)^{2s+1}}\right\},$$

$$\mathrm{II}_7=-\frac{(-1)^s(2s)!2^{2s+1}}{\pi^{2s+1}}$$

$$\cdot\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left[\frac{(-1)^m}{\{2m-1+(2n-1)ci\}^{2s+1}}+\frac{(-1)^m}{\{2m-1-(2n-1)ci\}^{2s+1}}\right],$$

$$\mathrm{I}_8=\frac{(-1)^s(2s)!2^{2s}}{\pi^{2s+1}}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\frac{(-1)^{m+n}}{(2m-1+2nci)^{2s+1}}+\frac{(-1)^{m+n}}{(2m-1-2nci)^{2s+1}}\right\},$$

$$\mathrm{II}_8=-\frac{i(-1)^s(2s-1)!2^{2s}}{\pi^{2s}}$$

$$\cdot\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left[\frac{(-1)^{m+n}}{\{2m-1+(2n-1)ci\}^{2s}}-\frac{(-1)^{m+n}}{\{2m-1-(2n-1)ci\}^{2s}}\right].$$

It is seen that the series in the two groups are each expanded into a double series, together with a single series in the first eight cases. The order of summation of the double series is interchangeable if the exponent $2s$ or $2s+1$ of the terms is not less than 3. It is no longer interchangeable if the exponent is 2 or 1. These expansions hold for any real c in general.

3. Summation of double series. The four types of single series involved in the first eight expansions of the series can be summed in closed form in terms of π by (13). To sum the double series, we define for any real c,

$$\sigma_{2s}^{*}(ci)\sum_{n,m=-\infty}^{\infty}{}'\frac{1}{(m+nci)^{2s}},\qquad s\geqq 2,$$

(16)

$$W_1(z|ci)=\frac{1}{z}+\sum_{n,m=-\infty}^{\infty}{}'\left\{\frac{1}{z-m-nci}+\frac{1}{m+nci}+\frac{z}{(m+nci)^2}\right\},$$

$$W_2(z|ci)=\frac{1}{z^2}+\sum_{n,m=-\infty}^{\infty}{}'\left\{\frac{1}{(z-m-nci)^2}-\frac{1}{(m+nci)^2}\right\},$$

$$W_s(z|ci)=\sum_{n,m=-\infty}^{\infty}\frac{1}{(z-m-nci)^s},\qquad s\geqq 3.$$

As before, the prime on the first three summation signs denotes the omission of simultaneous zeros of m and n. For $s\geqq 2$, W_s is an elliptic function of double periods 1 and ci and for $s=1$, W_1 is a pseudo-elliptic function [9]. W_1 and W_2 are known as Weierstrass zeta and elliptic functions, respectively. It appears that the double series involved in the expansions of the series can be expressed in terms of the coefficient σ_{2s}^* and the function W_s at half and quarter periods.

By decomposing the double series in σ_{2s}^*, we find for $s \geqq 2$,

$$\sigma_{2s}^*(ci) = 2\left\{1+\frac{(-1)^s}{c^{2s}}\right\} S_{2s} + 2\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\frac{1}{(m+nci)^{2s}}+\frac{1}{(m-nci)^{2s}}\right\}. \tag{17}$$

If this expression is employed to define σ_{2s}^*, its validity can be extended to include the case $s=1$ provided that in this case the order of summation of the double series is restricted to be not interchangeable. By similarly decomposing the double series in the function W_{2s} at half periods, we find for $s \geqq 1$:

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\frac{1}{(m+nci)^{2s}}+\frac{1}{(m-nci)^{2s}}\right\} = \frac{1}{2}\sigma_{2s}^*(ci) - \left\{1+\frac{(-1)^s}{c^{2s}}\right\} S_{2s},$$

$$\begin{aligned}
&\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\frac{1}{(2m-1+2nci)^{2s}}+\frac{1}{(2m-1-2nci)^{2s}}\right\} \\
&\quad = \frac{1}{2^{2s+1}} W_{2s}\left(\frac{1}{2}\middle| ci\right) - U_{2s} + \frac{1}{8}\delta_{1,s}\sigma_2^*(ci), \\
&\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left[\frac{1}{\{2m+(2n-1)ci\}^{2s}}+\frac{1}{\{2m-(2n-1)ci\}^{2s}}\right] \\
&\quad = \frac{1}{2^{2s+1}} W_{2s}\left(\frac{1}{2}ci\middle| ci\right) - \frac{(-1)^s}{c^{2s}} U_{2s} + \frac{1}{8}\delta_{1,s}\sigma_2^*(ci), \\
&\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left[\frac{1}{\{2m-1+(2n-1)ci\}^{2s}}+\frac{1}{\{2m-1-(2n-1)ci\}^{2s}}\right] \\
&\quad = \frac{1}{2^{2s+1}} W_{2s}\left(\frac{1}{2}+\frac{1}{2}ci\middle| ci\right) + \frac{1}{8}\delta_{1,s}\sigma_2^*(ci),
\end{aligned} \tag{18}$$

where $\delta_{n,s}$ is Kronecker delta. Note that the case $s=1$ is in general a special case to be considered separately. By decomposing the following double series into the difference of two double series of positive terms, we further find, for $s \geqq 1$,

$$\begin{aligned}
&\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\frac{(-1)^m}{(m+nci)^{2s}}+\frac{(-1)^m}{(m-nci)^{2s}}\right\} \\
&\quad = \frac{1}{2^{2s+1}}\left\{\sigma_{2s}^*\left(\frac{1}{2}ci\right) - W_{2s}\left(\frac{1}{2}\middle|\frac{1}{2}ci\right)\right\} + S_{2s}^* - \frac{(-1)^s}{c^{2s}} S_{2s} - \frac{1}{8}\delta_{1,s}\sigma_2^*\left(\frac{1}{2}ci\right), \\
&\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\frac{(-1)^n}{(m+nci)^{2s}}+\frac{(-1)^n}{(m-nci)^{2s}}\right\} \\
&\quad = \frac{1}{2}\sigma_{2s}^*(2ci) - \frac{1}{2}W_{2s}(ci|2ci) - S_{2s} + \frac{(-1)^s}{c^{2s}} S_{2s}^* - \frac{1}{2}\delta_{1,s}\sigma_2^*(2ci), \\
&\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left[\frac{(-1)^m}{\{2m+(2n-1)ci\}^{2s}}+\frac{(-1)^m}{\{2m-(2n-1)ci\}^{2s}}\right]
\end{aligned}$$

(19)
$$=\frac{1}{2^{4s+1}}\left\{W_{2s}\left(\frac{1}{4}ci\,\middle|\,\frac{1}{2}ci\right)-W_{2s}\left(\frac{1}{2}+\frac{1}{4}ci\,\middle|\,\frac{1}{2}ci\right)\right\}-\frac{(-1)^s}{c^{2s}}U_{2s},$$

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\frac{(-1)^n}{(2m-1+2nci)^{2s}}+\frac{(-1)^n}{(2m-1-2nci)^{2s}}\right\}$$
$$=\frac{1}{2^{2s+1}}\left\{W_{2s}\left(\frac{1}{2}\,\middle|\,2ci\right)-W_{2s}\left(\frac{1}{2}+ci\,\middle|\,2ci\right)\right\}-U_{2s}.$$

Again, by decomposing the following double series into four double series, we similarly find for $s\geqq 1$,

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\frac{(-1)^{m+n}}{(m+nci)^{2s}}+\frac{(-1)^{m+n}}{(m-nci)^{2s}}\right\}$$
$$=\frac{1}{2^{2s}}W_{2s}\left(\frac{1}{2}+\frac{1}{2}ci\,\middle|\,ci\right)-\left(\frac{1}{2}-\frac{1}{2^{2s}}\right)\sigma_{2s}^{*}(ci)$$

(20)
$$+\left\{1+\frac{(-1)^s}{c^{2s}}\right\}S_{2s}^{*}+\frac{1}{4}\delta_{1,s}\sigma_{2}^{*}(ci),$$

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left[\frac{(-1)^{m+n}}{\{2m-1+(2n-1)ci\}^{2s}}-\frac{(-1)^{m+n}}{\{2m-1-(2n-1)ci\}^{2s}}\right]$$
$$=\frac{1}{4^{2s}}\left\{W_{2s}\left(\frac{1}{4}+\frac{1}{4}ci\,\middle|\,ci\right)-W_{2s}\left(\frac{1}{4}-\frac{1}{4}ci\,\middle|\,ci\right)\right\}.$$

Note that the first expression has been simplified to the present form with the aid of the following relations:

(21)
$$W_2(\tfrac{1}{2}|ci)+W_2(\tfrac{1}{2}ci|ci)+W_2(\tfrac{1}{2}+\tfrac{1}{2}ci|ci)=0,$$
$$W_{2s}(\tfrac{1}{2}|ci)+W_{2s}(\tfrac{1}{2}ci|ci)+W_{2s}(\tfrac{1}{2}+\tfrac{1}{2}ci|ci)=(2^{2s}-1)\sigma_{2s}^{*}(ci),\qquad s\geqq 2.$$

Likewise, we further find for $s\geqq 0$,

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\frac{(-1)^m}{(2m-1+2nci)^{2s+1}}+\frac{(-1)^m}{(2m-1-2nci)^{2s+1}}\right\}$$
$$=-\frac{1}{4^{2s+1}}W_{2s+1}\left(\frac{1}{4}\,\middle|\,\frac{1}{2}ci\right)+U_{2s+1}^{*}+\frac{1}{16}\delta_{0,s}\sigma_{2}^{*}\left(\frac{1}{2}ci\right),$$

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left[\frac{(-1)^n}{\{2m+(2n-1)ci\}^{2s+1}}-\frac{(-1)^n}{\{2m-(2n-1)ci\}^{2s+1}}\right]$$
$$=-\frac{1}{2^{2s+1}}W_{2s+1}\left(\frac{1}{2}ci\,\middle|\,2ci\right)-\frac{i(-1)^s}{c^{2s+1}}U_{2s+1}^{*}+\frac{ci}{4}\delta_{0,s}\sigma_{2}^{*}(2ci),$$

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left[\frac{(-1)^m}{\{2m-1+(2n-1)ci\}^{2s+1}}+\frac{(-1)^m}{\{2m-1-(2n-1)ci\}^{2s+1}}\right]$$
$$=-\frac{1}{4^{2s+1}}W_{2s+1}\left(\frac{1}{4}+\frac{1}{4}ci\,\middle|\,\frac{1}{2}ci\right)+\frac{1+ci}{16}\delta_{0,s}\sigma_{2}^{*}\left(\frac{1}{2}ci\right),$$

(22)

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left[\frac{(-1)^n}{\{2m-1+(2n-1)ci\}^{2s+1}}-\frac{(-1)^n}{\{2m-1-(2n-1)ci\}^{2s+1}}\right]$$
$$=-\frac{1}{2^{2s+1}}W_{2s+1}\left(\frac{1}{2}+\frac{1}{2}ci\,\middle|\,2ci\right)+\frac{1+ci}{4}\delta_{0,s}\sigma_2^*(2ci),$$

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\frac{(-1)^{m+n}}{(2m-1+2nci)^{2s+1}}+\frac{(-1)^{m+n}}{(2m-1-2nci)^{2s+1}}\right\}$$
$$=\frac{1}{4^{2s+1}}\left\{W_{2s+1}\left(\frac{1}{4}+\frac{1}{2}ci\,\middle|\,ci\right)-W_{2s+1}\left(\frac{1}{4}\,\middle|\,ci\right)\right\}+U_{2s+1}^*-\frac{ci}{8}\delta_{0,s}\sigma_2^*(ci),$$

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left[\frac{(-1)^{m+n}}{\{2m+(2n-1)ci\}^{2s+1}}-\frac{(-1)^{m+n}}{\{2m-(2n-1)ci\}^{2s+1}}\right]$$
$$=\frac{1}{4^{2s+1}}\left\{W_{2s+1}\left(\frac{1}{2}+\frac{1}{4}ci\,\middle|\,ci\right)-W_{2s+1}\left(\frac{1}{4}ci\,\middle|\,ci\right)-\frac{i(-1)^s}{c^{2s+1}}U_{2s+1}^*\right.$$
$$\left.-\frac{1}{8}\delta_{0,s}\sigma_2^*(ci).\right.$$

Some of the summations of the double series were also found in the previous papers [7], [10].

4. The results. With the foregoing summations, the following results are obtained from (11), (12), (14) and (15):

$$\mathrm{I}_1=\frac{(-1)^s(2s-1)!}{2(2\pi)^{2s}}\{\sigma_{2s}^*(ci)-2S_{2s}\},$$

$$\mathrm{I}_2=\frac{(-1)^s(2s-1)!}{2(2\pi)^{2s}}\{W_{2s}(ci|2ci)-\sigma_{2s}^*(2ci)+2S_{2s}+\delta_{1,s}\sigma_2^*(2ci)\},$$

$$\mathrm{I}_3=\frac{(-1)^s(2s-1)!}{4(2\pi)^{2s}}\left\{\sigma_{2s}^*\left(\frac{1}{2}ci\right)\right.$$
$$\left.-W_{2s}\left(\frac{1}{2}\,\middle|\,\frac{1}{2}ci\right)+(2^{2s+1}-4)S_{2s}-\delta_{1,s}\sigma_2^*\left(\frac{1}{2}ci\right)\right\},$$

$$\mathrm{I}_4=-\frac{(-1)^s(2s-1)!}{2(2\pi)^{2s}}\left\{W_{2s}\left(\frac{1}{2}+\frac{1}{2}ci\,\middle|\,ci\right)\right.$$
$$\left.-(2^{2s-1}-1)\sigma_{2s}^*(ci)+2^{2s}S_{2s}^*+\delta_{1,s}\sigma_2^*(ci)\right\},$$

(23)

$$\mathrm{I}_5=-\frac{(-1)^s(2s-1)!}{2(2\pi)^{2s}}\left\{W_{2s}\left(\frac{1}{2}\,\middle|\,ci\right)-2^{2s+1}U_{2s}+\delta_{1,s}\sigma_2^*(ci)\right\},$$

$$\mathrm{I}_6=\frac{(-1)^s(2s)!}{2(2\pi)^{2s}}\left\{W_{2s}\left(\frac{1}{2}\,\middle|\,2ci\right)-W_{2s}\left(\frac{1}{2}+ci\,\middle|\,2ci\right)-2^{2s+1}U_{2s}\right\},$$

$$\mathrm{I}_7=\frac{(-1)^s(2s)!}{2(2\pi)^{2s+1}}\left\{W_{2s+1}\left(\frac{1}{4}\,\middle|\,\frac{1}{2}ci\right)-4^{2s+1}U_{2s+1}^*-\frac{1}{4}\delta_{0,s}\sigma_2^*\left(\frac{1}{2}ci\right)\right\},$$

$$\mathrm{I}_8=\frac{(-1)^s(2s)!}{2(2\pi)^{2s+1}}\left\{W_{2s+1}\left(\frac{1}{4}+\frac{1}{2}ci\,\middle|\,ci\right)\right.$$
$$\left.-W_{2s+1}\left(\frac{1}{4}\,\middle|\,ci\right)+4^{2s+1}U^*_{2s+1}+\frac{ci}{2}\delta_{0,s}\sigma_2^*(ci)\right\},$$

and

$$\mathrm{II}_1=\frac{(-1)^s(2s-1)!}{(2\pi)^{2s}}\left\{W_{2s}\left(\frac{1}{2}ci\,\middle|\,ci\right)+\delta_{1,s}\sigma_2^*(ci)\right\},$$

$$\mathrm{II}_2=\frac{2(-1)^s(2s)!}{(2\pi)^{2s+1}}\left\{iW_{2s+1}\left(\frac{1}{2}ci\,\middle|\,2ci\right)+\frac{c}{2}\delta_{0,s}\sigma_2^*(2ci)\right\},$$

$$\mathrm{II}_3=\frac{(-1)^s(2s-1)!}{2(2\pi)^{2s}}\left\{W_{2s}\left(\frac{1}{4}ci\,\middle|\,\frac{1}{2}ci\right)-W_{2s}\left(\frac{1}{2}+\frac{1}{4}ci\,\middle|\,\frac{1}{2}ci\right)\right\},$$

$$\mathrm{II}_4=\frac{i(-1)^s(2s)!}{(2\pi)^{2s+1}}\left\{W_{2s+1}\left(\frac{1}{4}ci\,\middle|\,ci\right)-W_{2s+1}\left(\frac{1}{2}+\frac{1}{4}ci\,\middle|\,ci\right)+\frac{1}{2}\delta_{0,s}\sigma_2^*(ci)\right\},$$

(24)

$$\mathrm{II}_5=-\frac{(-1)^s(2s-1)!}{(2\pi)^{2s}}\left\{W_{2s}\left(\frac{1}{2}+\frac{1}{2}ci\,\middle|\,ci\right)+\delta_{1,s}\sigma_2^*(ci)\right\},$$

$$\mathrm{II}_6=-\frac{2i(-1)^s(2s)!}{(2\pi)^{2s+1}}\left\{W_{2s+1}\left(\frac{1}{2}+\frac{1}{2}ci\,\middle|\,2ci\right)-\frac{1+ci}{2}\delta_{0,s}\sigma_2^*(2ci)\right\},$$

$$\mathrm{II}_7=\frac{(-1)^s(2s)!}{(2\pi)^{2s+1}}\left\{W_{2s+1}\left(\frac{1}{4}+\frac{1}{4}ci\,\middle|\,\frac{1}{2}ci\right)-\frac{1+ci}{4}\delta_{0,s}\sigma_2^*\left(\frac{1}{2}ci\right)\right\},$$

$$\mathrm{II}_8=-\frac{i(-1)^s(2s-1)!}{(2\pi)^{2s}}\left\{W_{2s}\left(\frac{1}{4}+\frac{1}{4}ci\,\middle|\,ci\right)-W_{2s}\left(\frac{1}{4}-\frac{1}{4}ci\,\middle|\,ci\right)\right\}.$$

It is seen that the two groups of series are thus expressed in terms of the coefficient σ^*_{2s} and the function W_s at half and quarter periods, of double periods $(1, ci)$, $(1, ci/2)$ or $(1, 2ci)$. The relations hold for any real value of c. In particular, when $c=1, \sqrt{3}$ or $1/\sqrt{3}$, the values of the coefficients σ_4^*, σ_6^* and the functions W_1, W_2, W_2' at half and quarter periods have been tabulated by the author in a recent paper [11] for the three preceding double periods corresponding to the three particular values of c. The values are expressed in closed form in terms of σ_4 when $c=1$ and in terms of σ_6 when $c=\sqrt{3}$ or $1/\sqrt{3}$. It is also noted that the values of σ_2^* and W_3 are

(25)
$$\sigma_2^*(ci)=2W_1(\tfrac{1}{2}|ci),$$
$$W_3(z|ci)=-\tfrac{1}{2}W_2'(z|ci).$$

Further values of σ^*_{2s} can be found successively in terms of σ_4^* and σ_6^* from the following recurrence relation, for $s\geqq 4$:

(26)
$$\tfrac{1}{3}(s-3)(2s+1)G_{2s}=G_4G_{2s-4}+G_6G_{2s-6}+\cdots+G_{2s-4}G_4,$$

where, for $s\geqq 2$,

(27)
$$G_{2s}=(2s-1)\sigma^*_{2s}(ci).$$

The value of W_4 is given by

$$W_4(z|ci) = W_2^2(z|ci) - 5\sigma_4^*(ci). \tag{28}$$

A similar recurrence relation for W_s with $s \geqq 5$ is

$$\tfrac{1}{6}(s-2)(s-3)F_s = F_2F_{s-2} + F_3F_{s-3} + \cdots + F_{s-2}F_2, \tag{29}$$

where, for $s \geqq 2$,

$$F_s = (s-1)W_s(z|ci). \tag{30}$$

Hence with the values of σ_{2s}^* and W_s so obtained, the two groups of series can be expressed in closed form in terms of σ_4 when $c = 1$ and in terms of σ_6 when $c = \sqrt{3}$ or $1/\sqrt{3}$. Note that some summations may not involve σ_4 or σ_6. The summations of each series for $s = 1$ and 2 are shown in Tables 1 and 2 (following § 5), where the following notations are used for shortness:

$$u = (15\sigma_4)^{1/2}, \qquad v = (35\sigma_6)^{1/3}. \tag{31}$$

The summations of the series I_7 and I_8 for $s = 0$ are shown in Table 3.

TABLE 3

c	I_7 $(s=0)$	I_8 $(s=0)$
1	$-\frac{1}{4}+\frac{\sqrt{2u}}{4\pi}$	$\frac{1}{4}-\frac{\sqrt{u}}{4\pi}$
$\sqrt{3}$	$-\frac{1}{4}+\frac{3^{1/4}\sqrt{v}}{4\pi}$	$\frac{1}{4}-\frac{3^{1/4}2^{1/2}(\sqrt{3}+1)\sqrt{v}}{16\pi}$
$1/\sqrt{3}$	$-\frac{1}{4}+\frac{3^{1/4}\sqrt{3v}}{4\pi}$	$\frac{1}{4}-\frac{3^{1/4}2^{1/2}(3-\sqrt{3})\sqrt{v}}{16\pi}$

Those of series II_2, II_4, II_6 and II_7 for $s = 0$ have been tabulated previously [10]. Note that when $s = 0$, $\mathrm{II}_4 = \mathrm{II}_7$.

The results confirm two of Ramanujan's summations in (3) namely, (i) and the particular case of (v) when $s = 1$. The summations (ii), (iii) and (iv) can be confirmed without difficulty. To confirm those of (v), (vi) and (vii) in general, it is necessary to show that in the expression of II_8 in (24), the following relations hold for $s \geqq 1$:

$$\begin{aligned} &W_{4s}(\tfrac{1}{4}+\tfrac{1}{4}ci|ci) = W_{4s}(\tfrac{1}{4}-\tfrac{1}{4}ci|ci) \quad \text{when } c = 1, \\ &W_{6s}(\tfrac{1}{4}+\tfrac{1}{4}ci|ci) = W_{6s}(\tfrac{1}{4}-\tfrac{1}{4}ci|ci) \quad \text{when } c = \sqrt{3} \text{ or } 1/\sqrt{3}. \end{aligned} \tag{32}$$

They can be proved readily by expanding the functions into series of their arguments. Each such function is real.

Besides, a summation analogous to the last three summations in (3) is given by Sandham [5] as follows:

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1} n^{4s+1}}{\sinh n\pi} = 0, \qquad s \geqq 1. \tag{33}$$

It is the series (iii) in (7), save $s = 0$. By using the expression II_5 in (24), the result follows at once on account of the following relation when $c = 1$:

$$W_{2s+1}(\tfrac{1}{2} + \tfrac{1}{2}ci \,|\, ci) = 0, \qquad s \geqq 0. \tag{34}$$

5. Conclusiveness of list. In the foregoing, it is seen that there are altogether eight kinds of operations which can be manipulated on the four equations in (10). On the left of the equations, a total of 32 series can be developed, which contain exponential or hyperbolic functions. Of them 16 are listed in (1) and (2). On the right, a single series and a double series are developed in each case from the first two equations. The single series can be summed in closed form in terms of π whenever it belongs to the four types in (13). This immediately excludes eight of the 16 single series. In the eight single series which can be so summed, it happens that the accompanying double series can be summed in closed form too in terms of σ_4 or σ_6 for the three particular values of c. The only exception is the particular case in which the single series is of the form

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}. \tag{35}$$

Although it can be summed to ln 2, yet the accompanying double series cannot be summed in closed form in the preceding manner. On the other hand, a double series only is developed on the right in each case from the last two equations. It is found that besides the eight double series in (14) and (15), none of the other eight double series can be so summed. These investigations lead us to conclude that the list of the 16 series shown in (1) and (2) is indeed conclusive of its kind.

TABLE 1

Summations of series in Group I for s = 1 and 2

Series	c	s = 1	s = 2
I_1	1	$\frac{1}{24}-\frac{1}{8\pi}$	$-\frac{1}{240}+\frac{u^2}{80\pi^4}$
	$\sqrt{3}$	$\frac{1}{24}-\frac{\sqrt{3}}{24\pi}-\frac{v}{32\pi^2}$	$-\frac{1}{240}+\frac{3v^2}{256\pi^4}$
	$1/\sqrt{3}$	$\frac{1}{24}-\frac{\sqrt{3}}{8\pi}+\frac{3v}{32\pi^2}$	$-\frac{1}{240}+\frac{27v^2}{256\pi^4}$
I_2	1	$-\frac{1}{24}+\frac{u}{16\pi^2}$	$\frac{1}{240}-\frac{3u^2}{640\pi^4}$
	$\sqrt{3}$	$-\frac{1}{24}+\frac{(2\sqrt{3}+1)v}{64\pi^2}$	$\frac{1}{240}-\frac{3(4\sqrt{3}+1)v^2}{2048\pi^4}$

TABLE 1 (cont.)

Series	c	$s=1$	$s=2$
	$1/\sqrt{3}$	$-\frac{1}{24}+\frac{3(2\sqrt{3}-1)v}{64\pi^2}$	$\frac{1}{240}+\frac{27(4\sqrt{3}-1)v^2}{2048\pi^4}$
I_3	1	$-\frac{1}{24}+\frac{u}{8\pi^2}$	$\frac{7}{240}+\frac{3u^2}{80\pi^4}$
	$\sqrt{3}$	$-\frac{1}{24}+\frac{(2\sqrt{3}-1)v}{32\pi^2}$	$\frac{7}{240}-\frac{3(4\sqrt{3}-1)v^2}{256\pi^4}$
	$1/\sqrt{3}$	$-\frac{1}{24}+\frac{3(2\sqrt{3}+1)v}{32\pi^2}$	$\frac{7}{240}+\frac{27(4\sqrt{3}+1)v^2}{256\pi^4}$
I_4	1	$\frac{1}{24}$	$-\frac{7}{240}+\frac{3u^2}{20\pi^4}$
	$\sqrt{3}$	$\frac{1}{24}-\frac{v}{16\pi^2}$	$-\frac{7}{240}+\frac{3v^2}{32\pi^4}$
	$1/\sqrt{3}$	$\frac{1}{24}+\frac{3v}{16\pi^2}$	$-\frac{7}{240}+\frac{27v^2}{32\pi^4}$
I_5	1	$-\frac{1}{8}+\frac{1}{8\pi}+\frac{u}{8\pi^2}$	$\frac{1}{16}-\frac{u^2}{8\pi^4}$
	$\sqrt{3}$	$-\frac{1}{8}+\frac{\sqrt{3}}{24\pi}+\frac{(\sqrt{3}+1)v}{16\pi^2}$	$\frac{1}{16}-\frac{3(2+\sqrt{3})v^2}{64\pi^4}$
	$1/\sqrt{3}$	$-\frac{1}{8}+\frac{\sqrt{3}}{8\pi}+\frac{3(\sqrt{3}-1)v}{16\pi^2}$	$\frac{1}{16}-\frac{27(2-\sqrt{3})v^2}{64\pi^4}$
I_6	1	$\frac{1}{8}-\frac{\sqrt{2}u}{8\pi^2}$	$-\frac{1}{16}+\frac{3\sqrt{2}u^2}{32\pi^4}$
	$\sqrt{3}$	$\frac{1}{8}-\frac{\sqrt{2}(3+\sqrt{3})v}{32\pi^2}$	$-\frac{1}{16}+\frac{3\sqrt{2}(7\sqrt{3}+9)v^2}{512\pi^4}$
	$1/\sqrt{3}$	$\frac{1}{8}-\frac{3\sqrt{2}(3-\sqrt{3})v}{32\pi^2}$	$-\frac{1}{16}+\frac{27\sqrt{2}(7\sqrt{3}-9)v^2}{512\pi^4}$
I_7	1	$\frac{1}{4}-\frac{\sqrt{2}}{4\pi^3}u^{3/2}$	$-\frac{5}{4}+\frac{9\sqrt{2}}{4\pi^5}u^{5/2}$
	$\sqrt{3}$	$\frac{1}{4}-\frac{3^{1/4}(3+2\sqrt{3})}{16\pi^3}v^{3/2}$	$-\frac{5}{4}+\frac{3\cdot 3^{1/4}(39+20\sqrt{3})}{64\pi^5}v^{5/2}$

TABLE 1 (cont.)

Series	c	$s=1$	$s=2$
	$1/\sqrt{3}$	$\frac{1}{4}-\frac{9\cdot 3^{1/4}(2-\sqrt{3})}{16\pi^3}v^{3/2}$	$-\frac{5}{4}+\frac{81\cdot 3^{1/4}(13\sqrt{3}-20)}{64\pi^5}v^{5/2}$
I_8	1	$-\frac{1}{4}+\frac{1}{2\pi^3}u^{3/2}$	$\frac{5}{4}-\frac{3}{\pi^5}u^{5/2}$
	$\sqrt{3}$	$-\frac{1}{4}+\frac{3^{1/4}2^{1/2}(3+\sqrt{3})}{16\pi^3}v^{3/2}$	$\frac{5}{4}-\frac{3\cdot 3^{1/4}2^{1/2}(3+2\sqrt{3})}{8\pi^5}v^{5/2}$
	$1/\sqrt{3}$	$-\frac{1}{4}+\frac{9\cdot 3^{1/4}2^{1/2}(\sqrt{3}-1)}{16\pi^3}v^{3/2}$	$\frac{5}{4}-\frac{81\cdot 3^{1/4}2^{1/2}(2-\sqrt{3})}{8\pi^5}v^{5/2}$

TABLE 2
Summations of series in Group II *for s* = 1 *and* 2

Series	c	$s=1$	$s=2$
II_1	1	$-\frac{1}{4\pi}+\frac{u}{4\pi^2}$	$\frac{u^2}{4\pi^4}$
	$\sqrt{3}$	$-\frac{\sqrt{3}}{12\pi}+\frac{(\sqrt{3}-1)v}{8\pi^2}$	$\frac{3(2-\sqrt{3})v^2}{32\pi^4}$
	$1/\sqrt{3}$	$-\frac{\sqrt{3}}{4\pi}+\frac{3(\sqrt{3}+1)v}{8\pi^2}$	$\frac{27(2+\sqrt{3})v^2}{32\pi^4}$
II_2	1	$\frac{\sqrt{2}}{8\pi^3}u^{3/2}$	$\frac{9\sqrt{2}}{32\pi^5}u^{5/2}$
	$\sqrt{3}$	$\frac{3^{1/4}(2\sqrt{3}-3)}{32\pi^3}u^{3/2}$	$\frac{3\cdot 3^{1/4}(39-20\sqrt{3})}{512\pi^5}v^{5/2}$
	$1/\sqrt{3}$	$\frac{9\cdot 3^{1/4}(2+\sqrt{3})}{32\pi^3}v^{3/2}$	$\frac{81\cdot 3^{1/4}(13\sqrt{3}+20)}{512\pi^5}v^{5/2}$
II_3	1	$\frac{\sqrt{2}u}{2\pi^2}$	$\frac{3\sqrt{2}u^2}{2\pi^4}$
	$\sqrt{3}$	$\frac{\sqrt{2}(3-\sqrt{3})v}{8\pi^2}$	$\frac{3\sqrt{2}(7\sqrt{3}-9)v^2}{32\pi^4}$
	$1/\sqrt{3}$	$\frac{3\sqrt{2}(3+\sqrt{3})v}{8\pi^2}$	$\frac{27\sqrt{2}(7\sqrt{3}+9)v^2}{32\pi^4}$

TABLE 2 (cont.)

Series	c	$s=1$	$s=2$
II_4	1	$\frac{1}{\pi^3}u^{3/2}$	$\frac{6}{\pi^5}u^{5/2}$
	$\sqrt{3}$	$\frac{2^{1/2}3^{1/4}(3-\sqrt{3})}{8\pi^3}v^{3/2}$	$\frac{3\cdot 2^{1/2}3^{1/4}(2\sqrt{3}-3)}{4\pi^5}v^{5/2}$
	$1/\sqrt{3}$	$\frac{9\cdot 2^{1/2}3^{1/4}(\sqrt{3}+1)}{8\pi^3}v^{3/2}$	$\frac{81\cdot 2^{1/2}3^{1/4}(2+\sqrt{3})}{4\pi^5}v^{5/2}$
II_5	1	$\frac{1}{4\pi}$	$\frac{u^2}{8\pi^4}$
	$\sqrt{3}$	$\frac{\sqrt{3}}{12\pi}-\frac{v}{16\pi^2}$	$\frac{3v^2}{128\pi^4}$
	$1/\sqrt{3}$	$\frac{\sqrt{3}}{4\pi}+\frac{3v}{16\pi^2}$	$\frac{27v^2}{128\pi^4}$
II_6	1	$\frac{1}{8\pi^3}u^{3/2}$	$\frac{3}{32\pi^5}u^{5/2}$
	$\sqrt{3}$	$\frac{3^{1/4}2^{1/2}(3-\sqrt{3})}{128\pi^3}v^{3/2}$	$\frac{3\cdot 3^{1/4}2^{1/2}(9+\sqrt{3})}{2048\pi^5}v^{5/2}$
	$1/\sqrt{3}$	$\frac{9\cdot 3^{1/4}2^{1/2}(\sqrt{3}+1)}{128\pi^3}v^{3/2}$	$-\frac{81\cdot 3^{1/4}2^{1/2}(3\sqrt{3}-1)}{2048\pi^5}v^{5/2}$
II_7	1	$\frac{1}{2\pi^3}u^{3/2}$	$-\frac{3}{2\pi^5}u^{5/2}$
	$\sqrt{3}$	$\frac{3^{1/4}2^{1/2}(3+\sqrt{3})}{32\pi^3}v^{3/2}$	$\frac{3\cdot 3^{1/4}2^{1/2}(9-\sqrt{3})}{128\pi^5}v^{5/2}$
	$1/\sqrt{3}y$	$\frac{9\cdot 3^{1/4}2^{1/2}(\sqrt{3}-1)}{32\pi^3}v^{3/2}$	$-\frac{81\cdot 3^{1/4}2^{1/2}(3\sqrt{3}-1)}{128\pi^5}v^{5/2}$
II_8	1	$\frac{u}{2\pi^2}$	0
	$\sqrt{3}$	$\frac{\sqrt{3}v}{8\pi^2}$	$\frac{3\sqrt{3}v^2}{16\pi^4}$
	$1/\sqrt{3}$	$\frac{3\sqrt{3}v}{8\pi^2}$	$-\frac{27\sqrt{3}v^2}{16\pi^4}$

REFERENCES

[1] S. Ramanujan, *Notebooks of Srinivasa Ramanujan*, 2 vols., Tata Institute, India, 1957.
[2] G. H. Hardy, P. V. Seshu Aiyar and B. H. Wilson, *Collected Papers of Srinivasa Ramanujan*, Cambridge University Press, Cambridge, 1927.
[3] G. N. Watson, *Theorems stated by Ramanujan (II): Theorems on summation of series*, J. London Math. Soc., 3 (1928), pp. 216–225.
[4] H. F. Sandham, *Three summations due to Ramanujan*, Quart. J. Math. Oxford Ser., 1 (1950), pp. 238–240.
[5] ———, *Some infinite series*, Proc. Lond. Math. Soc., 5 (1954), pp. 430–436.
[6] G. H. Hardy, *A formula of Ramanujan*, J. London Math. Soc., 3 (1928), pp. 238–240.
[7] C. B. Ling, *On summation of series of hyperbolic functions*, this Journal, 5 (1974), pp. 551–562.
[8] I. M. Ryshik and I. S. Gradstein, *Tables of Series, Products and Integrals*, 2nd ed., Plenum Press, New York, 1963, p. 36.
[9] E. T. Copson, *Theory of Functions of a Complex Variable*, Oxford University Press, Oxford, 1935.
[10] C. B. Ling, *On summation of series of hyperbolic functions. II*, this Journal, 6 (1957), pp. 129–139.
[11] ———, *On a class of Weierstrass elliptic functions at half and quarter periods*, this Journal, 6 (1975), pp. 117–128.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

ON THE LINEAR THEORY OF HEAT CONDUCTION FOR MATERIALS WITH MEMORY*

PAUL L. DAVIS†

Abstract. The linear theory of rigid conductors of heat composed of materials with memory is analyzed under assumptions consistent with the theory of Coleman and Gurtin. Under these assumptions, the resulting integro-differential equation is shown to be *parabolic modulo a trivial hyperbolic part*. An existence and uniqueness theorem follows.

1. Introduction. The linearized theory of rigid conductors of heat composed of materials with memory is based [1], [5], [9] on the constitutive relations

$$e = c\theta + \int_{-\infty}^{t} \alpha(t-s)\theta(s)\, ds$$

and

$$\mathbf{q} = -\kappa \nabla\theta - \int_{-\infty}^{t} \beta(t-s)\nabla\theta(s)\, ds$$

where θ denotes the departure of the temperature from its reference value, e the internal energy and $\mathbf{q}$ the heat flux. Together with the energy-balance law for a rigid stationary heat conductor,

$$\dot{e} = -\nabla \cdot \mathbf{q} + r,$$

where r is the heat supply, these relations imply

$$\frac{\partial}{\partial t}\left(c\theta(x,t) + \int_{-\infty}^{t} \alpha(t-s)\theta(x,s)\, ds\right) = \kappa\, \Delta\theta(x,t) + \int_{-\infty}^{t} \beta(t-s)\, \Delta\theta(x,s)\, ds + r(x,t) \tag{1.1}$$

where $x = (x_1, x_2, x_3)$.

When $\kappa = 0$, (1.1) is, under appropriate assumptions on α and β, a hyperbolic equation in the sense that signals propagate with finite speed [3], [4], [5]. Little is known about (1.1) when $\kappa \neq 0$; uniqueness theorems are contained in [8] and [9]. It is believed by many that (1) with $\kappa \neq 0$ is, in some sense, parabolic. This is almost true. We show in § 3 that for an appropriate class of kernels and $c\kappa > 0$, it is always *parabolic modulo a trivial hyperbolic part*. These notions are defined in § 2. They are intrinsic classification definitions [6]; that is, a definition that classifies an equation in accordance with the behavior of solutions of a certain type of problem. Consequently, an existence and uniqueness theorem is a consequence of classification. It is discussed in § 3.

2. Background. For convenience we assume throughout that $\theta(x,t) = 0$ for $t < t_1 \leqq 0$ and that θ is known for $t_1 \leqq t \leqq 0$. Integrating (1.1) and isolating all

* Received by the editors March 15, 1976.

† Department of Mathematics, Manhattanville College, Purchase, New York 10577.

known quantities, we have an equation of the form

$$
\begin{aligned}
\theta(x,t) = -c^{-1}\int_0^t \alpha(t-s)\theta(x,s)\,ds \\
+c^{-1}\int_0^t \left(\kappa + \int_0^{t-s} \beta(\tau)\,d\tau\right)\Delta\theta(x,s)\,ds + g(x,t).
\end{aligned}
\tag{2.1}
$$

This is a special case of the equation

$$
u(x,t) = \sum_{\nu=1}^{p} \int_0^t k_\nu(t-\tau)L_\nu u(x,\tau)\,d\tau + g(x,t) \tag{2.2}
$$

when $x = (x_1, \cdots, x_n)$ and L_ν is a constant coefficient differential operator with respect to these variables.

Let X be the space of continuous maps $g(x,t)$ from $[0,\infty)$ to $L^2(R^n)$ and let $Y \subset X$ be the subspace of C^∞ maps from $(0,\infty)$ to $L^2(R^n) \cap C^\infty(R^n)$. Let $Z \subset Y$ be the subspace of C^∞ maps from $(0,\infty)$ to $C_0^\infty(R^n)$. Let Z_1 be those elements of Z which have their support in $\{x \mid -b_i \leqq x_i \leqq b_i\}$ for each $t \geqq 0$.

DEFINITION. Equation (2.2) is *parabolic modulo a hyperbolic part* if

(a) for each $g \in Z_1$ there is a unique solution $u = S(g)$ of (2.2) contained in Y and

(b) there is nonzero $T: X \to Y$ such that

(i) $T(X) \cap Z = \{0\}$

and

(ii) for each $g \in Z_1$, $S(g) - T(g) \in Z$.

$S - T$ is called the hyperbolic part of the solution.

DEFINITION. Equation (2.2) is *parabolic* if the hyperbolic part is identically zero.

DEFINITION. The hyperbolic part is *trivial* if it maps Z_1 and Z_1.

Example. Equation (2.2) being parabolic modulo a trivial hyperbolic part means the nonparabolic part propagates with zero velocity. An elementary example of this phenomenon is the Cauchy problem for the partial differential equation

$$
\left(\frac{\partial}{\partial t} - \Delta\right)\frac{\partial^n}{\partial t^n}u = 0
$$

which can be reduced to (2.2) by integrating $(n+1)$ times with respect to t.

We investigate (1.1) for α and β in the class $\mathscr{A}$ consisting of those functions having rational Laplace transforms $\tilde{\alpha}(w)$ and $\tilde{\beta}(w)$ (degree of numerator less than that of denominator). Following [3], we have the following formula for the Fourier transform $\hat{\theta}(\xi, t)$ of the solution of (2.1) with respect to x:

$$
\hat{\theta}(\xi,t) = \hat{g}(\xi,t) + \int_0^t M(t-\tau,\xi)\hat{g}(\xi,\tau)\,d\tau, \tag{2.3}
$$

where

$$\tilde{M}(w,\xi)=\frac{-c^{-1}(w\tilde{\alpha}(w)+\kappa|\xi|^2+\tilde{\beta}(w)|\xi|^2)}{w+c^{-1}(w\tilde{\alpha}(w)+\kappa|\xi|^2+\tilde{\beta}(w)|\xi|^2)}. \tag{2.4}$$

Proceeding as in [2] or [3], we are led to the study of the singularities of $\tilde{M}(w,\xi)$. Indeed, $\theta(\cdot\,,t)$ is in $L^2(R^n)$ if the singularities $w_i(\xi)$ of $\tilde{M}(w,\xi)$ have bounded real part for all real ξ. (Since α and β are in $\mathcal{A}$, $\tilde{M}$ has only a finite number of poles.)

Assume at least one of α and β is not identically zero. Let

$$\tilde{\alpha}(w)=\sum_{j=0}^{q-1}a_jw^j\Big/\sum_{j=0}^{q}b_jw^j \tag{2.5}$$

and

$$\tilde{\beta}(w)=\sum_{j=0}^{m-1}c_jw^j\Big/\sum_{j=0}^{m}d_jw^j \tag{2.6}$$

where b_q and d_m are not zero. (Assume that common factors are divided out.) The singularities of $\tilde{M}$ are the zeros of the polynomial

$$\begin{aligned}P(w,|\xi|)=&\sum_{k=0}^{m}b_qd_kw^{q+k+1}+\sum_{j=0}^{q-1}\sum_{k=0}^{m}(b_jd_k+c^{-1}a_jd_k)w^{j+k+1}\\&+|\xi|^2c^{-1}\kappa\sum_{j=0}^{q}\sum_{k=0}^{m-1}b_jd_kw^{j+k}+|\xi|^2c^{-1}\sum_{j=0}^{q}\sum_{k=0}^{m-1}c_kb_jw^{j+k}.\end{aligned} \tag{2.7}$$

The behavior of the roots $w(|\xi|)$ of $P(w,|\xi|)=0$ as $|\xi|\to\infty$ can be investigated by examining the roots of

$$R(\omega,z)=\omega^{m+q+1}z^2P\left(\frac{1}{\omega},\frac{1}{z}\right) \tag{2.8}$$

as $|z|\to 0$.

3. Analysis. We examine the polynomial

$$\begin{aligned}R(\omega,z)=&z^2\sum_{k=0}^{m}b_qd_k\omega^{m-k}+z^2\sum_{j=0}^{q-1}\sum_{k=0}^{m}(b_jd_k+c^{-1}a_jd_k)\omega^{m+q-k-j}\\&+c^{-1}\kappa\sum_{j=0}^{q}\sum_{k=0}^{m}b_jd_k\omega^{m+q+1-k-j}+c^{-1}\sum_{j=0}^{q}\sum_{k=0}^{m-1}c_kb_j\omega^{m+q+1-k-j}\end{aligned} \tag{3.1}$$

given by (2.8). The roots $(\omega_1(z),\cdots,\omega_l(z))$, $2\leqq l\leqq m+q+1$, satisfy

THEOREM 3.1. *Only one root approaches zero as $z\to 0$. That root can be written as*

$$\omega(z)=-c\kappa^{-1}z^2+\sum_{k=1}^{\infty}c_kz^{k+2}. \tag{3.2}$$

Proof. Since $R(\omega,0)$ is a polynomial with $c^{-1}\kappa b_qd_m\omega$ as lowest order term, the first statement is a consequence of $c^{-1}\kappa b_qd_m\neq 0$. The series (3.2) can be computed by constructing a Newton's diagram [7] for (3.1). We just present the analysis determined by the diagram. Let $z=t$ and $\omega=t^2u$. Consider

$$R(t^2u,t)=b_qd_mt^2(1+c^{-1}\kappa u)+t^2G_1(t,u)$$

and

$$G(t,u)=R(t^2u,t)/t^2.$$

It satisfies $G(0,-c\kappa^{-1})=0$ and $G_u(0,-c\kappa^{-1})\neq 0$. By the implicit function theorem, there is $u(t)$, analytic in a neighborhood of the origin such that $u(0)=-c\kappa^{-1}$ and $G(t,u(t))=0$. Letting $\omega(z)=z^2u(z)$, we have $R(\omega(z),z)=0$ where $\omega(z)$ satisfies (3.2).

An immediate consequence is

COROLLARY 3.1. *Only one of the roots of* (2.7), *given by*

$$(w_1(|\xi|),\cdots,w_l(|\xi|))=\left(\frac{1}{\omega_1(1/|\xi|)},\cdots,\frac{1}{\omega_l(1/|\xi|)}\right)$$

satisfies $|w(|\xi|)|\to\infty$ *as* $|\xi|\to\infty$. *Moreover, that root can be written as*

$$w_1(|\xi|)=-|\xi|^2\Big/\left(c\kappa^{-1}-\sum_{k=1}^{\infty}c_k/|\xi|^k\right). \tag{3.3}$$

We can now exhibit the Fourier transform of Tg and the trivial hyperbolic part using (2.3) and the representation of $M(t,\xi)$ given in [2]; namely

$$M(t,\xi)=\frac{1}{2\pi i}\sum_{j=1}^{l}e^{w_j(\xi)t}\int_{C(0,\delta(\xi,j))}e^{tw}\tilde{M}(w+w_j(\xi),\xi)\,dw$$

where $C(0,\delta)$ is a sufficiently small circle of radius δ about 0. The Fourier transform of Tg is

$$\int_0^t M_1(t-\tau,\xi)\hat{g}(\xi,\tau)\,d\tau \tag{3.4}$$

where

$$M_1(t,\xi)=\frac{1}{2\pi i}e^{w_1(\xi)t}\int_{C(0,\,\delta(\xi,\,1))}e^{tw}\tilde{M}(w+w_1(\xi),\xi)\,dw. \tag{3.5}$$

The remainder of (2.3) is the Fourier transform of the trivial hyperbolic part; i.e.

$$\hat{g}(\xi,t)+\int_0^t M_2(t-\tau,\xi)\hat{g}(\xi,\tau)\,d\tau \tag{3.6}$$

where $M_2=M-M_1$. The hyperbolic part is not zero, since if it were, we would have (3.6) equal to zero. We could solve by successive approximations and deduce that $\hat{g}$, and hence g, is zero. An application of the Paley Weiner theorem, as in [2], together with the first part of Corollary 3.1, implies that (3.5) defines a hyperbolic part that is trivial. Standard arguments imply that the conditions of the definition of parabolic modulo a hyperbolic part are not satisfied. We summarize:

THEOREM 3.2. *If α and β are in $\mathscr{A}$ and are not both zero, then* (2.1) *is parabolic modulo a trivial hyperbolic part. Therefore, within the class X, there exists a unique solution of* (2.1) *for each $g \in Z_1$.*

REFERENCES

[1] B. D. COLEMAN AND M. E. GURTIN, *Equipresence and constitutive equations for rigid heat conductors*, Z. Angew. Math. Phys., 18 (1967), pp. 199–208.

[2] P. L. DAVIS, *Hyperbolic integrodifferential equations*, Proc. Amer. Math. Soc., 47 (1975), pp. 155–160.

[3] ———, *On the hyperbolicity of the equations of the linear theory of heat conduction for materials with memory*, SIAM J. Appl. Math., 30 (1976), pp. 75–80.

[4] J. M. FINN AND L. T. WHEELER, *Wave propagational aspects of the generalized theory of heat conduction*, Z. Angew. Math. Phys., 23 (1972), pp. 927–940.

[5] M. E. GURTIN AND A. C. PIPKIN, *A general theory of heat conduction with finite wave speeds*, Arch. Rational Mech. Anal., 31 (1968), pp. 113–126.

[6] R. HERSH, *How to classify differential polynomials*, Amer. Math. Monthly, 80 (1973), pp. 641–654.

[7] E. HILLE, *Analytic Function Theory*, vol. II, Ginn, Boston, 1962.

[8] R. R. NACHLINGER AND L. WHEELER, *A uniqueness theorem for rigid heat conductors with memory*, Quart. Appl. Math., 31 (1973), pp. 267–273.

[9] J. W. NUNZIATO, *On heat conduction in materials with memory*, Quart. Appl. Math., 29 (1971), pp. 187–204.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

SATURATION THEOREMS CONNECTED WITH THE ABSTRACT WAVE EQUATION*

JOHN W. DETTMAN†

Abstract. The study of certain well-posed Cauchy problems for the abstract heat equation leads to the theory of C_0 semi-groups of operators. The relevant saturation theory for the semi-group as a strong approximation process leads to many important results in approximation theory and differential equations. In this paper, we consider a certain class of well-posed Cauchy problems for the abstract wave equation and the solution of them as strong approximation processes for either the initial values of the solution or its derivative. The saturation order for each of these processes is found to be t^2 and the saturation class is characterized in each case.

1. Introduction. Let X be a Banach space. Consider the abstract Cauchy problem for the heat equation in X: $u'(t)=Au(t)$, $u(0)=f$, where A is a closed linear operator, densely defined, and with nonvoid resolvent set. It is well-known that this problem is uniformly well-posed in $\bar{R}_+=\{t|0\leqq t<\infty\}$ (see [11]) if and only if A is the infinitesimal generator of a C_0 semi-group $\Omega(t)$ and the solution is $u(t)=\Omega(t)f, f\in D(A)$. The semi-group $\Omega(t)$ is a commutative strong approximation process since $\|\Omega(t)f-f\|\to 0$ as $t\to 0^+$ for each $f\in X$. The process is saturated with order t and the Favard class (saturation class) is $\widetilde{D(A)}^X$ (the relative completion of $D(A)$ in X) (see [7], [8]). These results are the basis for a wide variety of saturation theorems in approximation theory. From the point of view of differential equations the saturation theory gives information on the boundary behavior of the solution, i.e., at what rate does the solution approach the initial conditions.

The situation with the wave equation is not quite so clear. For one thing, the well-posed Cauchy problem for the abstract wave equation $u''(t)=Au(t)$ has not been completely characterized within the context of semi-group theory. Actually, Fattorini [11], DaPrato–Guisti [9], and Sova [17] have shown that the Cauchy problem for this equation is uniformly well-posed if and only if A generates a strongly continuous cosine function. They also give necessary and sufficient conditions on the resolvent of A for it to be the generator of a strongly continuous cosine function. However, this approach precludes our use of the results of semi-group theory in the study of the saturation problem.

In this paper, we shall consider the following problem $u''(t)=Au(t)$, $u(0)=\phi$, $u'(0)=\psi$. This problem is uniformly well-posed in $R=\{t|-\infty<t<\infty\}$ if $A=B^2+c^2I$, where B is the infinitesimal generator of a C_0 group and c^2 is a nonnegative constant. Conversely, in many important cases (e.g. if $X=L_p$, $1<p<\infty$, or X is a Hilbert space and A is a self-adjoint operator [11]) when the Cauchy problem is uniformly well-posed A is necessarily of the form B^2+c^2I where B is a group generator. We discuss another case of this in § 2. In § 3 we consider the relevant saturation theorems for the solution operators when $A=B^2$, the square of a C_0 group generator. In § 4 we deal with the corresponding

* Received by the editors February 19, 1976, and in revised form July, 28, 1976.
† Department of Mathematical Sciences, Oakland University, Rochester, Michigan 48063.

theorems in the case where $A = B^2 + c^2 I$. This depends heavily on a representation of the solution operators in the case $c^2 \neq 0$ in terms of the corresponding operators in the case $c = 0$. These results have been derived using the methods of related differential equations [4]–[6], [10]. Finally, in § 5 we analyze the orders of $\|u(t)-\phi\|$, $\|u'(t)-\psi\|$, and $\|t^{-1}[u(t)-\phi]-\psi\|$ as $t \to 0$.

Throughout this paper a very important concept will be that of *relative completion*. Let X be a Banach space and Y a proper normalized Banach subspace continuously imbedded in X. Let $S_Y(\rho) = \{f \in X \mid \|f\|_Y \leqq \rho\}$; then the completion of Y relative to X, denoted by $\tilde{Y}^X$, is defined by

$$\tilde{Y}^X = \bigcup_{\rho > 0} \overline{S_Y(\rho)}^X$$

In other words, $\tilde{Y}^X$ is the set of all elements $f \in X$ which are in the closure in X of some bounded sphere in Y. This concept was first introduced by Gagliardo [13]. It is developed further by Aronszajn and Gagliardo [1] where it is proved that $\tilde{Y}^X$ is a normalized Banach subspace of X under the norm

$$\|f\|_{\tilde{Y}^X} = \inf\{\rho > 0 \mid f \in \overline{S_Y(\rho)}^X\}.$$

It is also shown that if Y is reflexive then Y and $\tilde{Y}^X$ are equal with equal norms. The concept of relative completion has been used by Berens [3] and by Shapiro [16]. (See also Butzer and Nessel [8].)

2. A well-posed Cauchy problem. Throughout this paper we shall take the underlying space X to be a Banach space. We shall consider the abstract Cauchy problem for the wave equation: $u''(t) = Au(t)$, $u(0) = \phi$, $u'(0) = \psi$, where A is a closed linear operator with nonvoid resolvent set defined on a domain $D(A)$ dense in X. We shall say, following Fattorini, that the problem is uniformly well-posed in R if there exists a dense subspace D such that when $\phi, \psi \in D$ there exists a unique solution depending continuously on the data. In this case, there exists a strongly continuous bounded linear operator $\tilde{S}(t)$ such that the solution can be expressed as

$$u(t) = \tilde{S}(t)\phi + \tilde{T}(t)\psi,$$

where $\tilde{T}(t) = \int_0^t \tilde{S}(\tau)\, d\tau$. This equation is interpreted in the strong operator topology. Furthermore, $\tilde{S}(t)$ is of type $\leqq \omega$, i.e. there exists a positive constant M such that $\|\tilde{S}(t)\| \leqq M e^{\omega|t|}$ for some real number ω. The operator A is called the infinitesimal generator of $\tilde{S}(t)$.

For the moment, let us assume that the Cauchy problem is uniformly well-posed if $\phi \in D(A)$ and $\psi = 0$. Then it can be shown that

$$v(t) = \frac{1}{\sqrt{\pi t}} \int_0^\infty e^{-s^2/(4t)} \tilde{S}(s)\phi\, ds$$

solves the abstract Cauchy problem for the heat equation (see [4], [11], [14])

$$v'(t) = Av(t),$$

$$v(0) = \phi.$$

This means that A is the infinitesimal generator of a C_0 semi-group $\Omega(t)$, holomorphic in the right half-plane Re $(t)>0$. If $\tilde{S}(t)$ is of type $\leqq \omega$ then $\Omega(t)$ is of type $\leqq \omega^2$. Therefore, for $c^2 \geqq \omega^2$, there is a positive constant K such that

$$\|\Omega(t)\, e^{-c^2 t}\| \leqq K$$

and $\Omega(t)\, e^{-c^2 t}$ is an equibounded semi-group with infinitesimal generator $A - c^2 I$. This means that we can take the square root of $c^2 I - A$ (see [2]) and define

$$B = i(c^2 I - A)^{1/2}$$

which will be a closed linear operator densely defined in X. Therefore, $B^2 = A - c^2 I$ and $A = B^2 + c^2 I$. This shows the existence of B and the desired representation $A = B^2 + c^2 I$ but does not yet show that B is the infinitesimal generator of a C_0 group.

We have seen, having assumed that the Cauchy problem for the wave equation is uniformly well-posed for $\phi \in D(A)$ and $\psi = 0$, that A is a semi-group generator and $A = B^2 + c^2 I$. We now assume further that the Cauchy problem has a unique solution if $\phi \in D(A)$ and $\psi \in D(B) \cap R(B)$ [$R(B)$ is the range of B]. By Theorem 5.9 and Lemma 6.1 of [11], the Cauchy problem $u''(t) = (A - c^2 I)u(t) = B^2 u(t)$ has a unique solution for $u(0) = \phi$ and $u'(0) = B\phi$ or $-B\phi$. But then by Theorem 23.9.5 of [15] B is the infinitesimal generator of a C_0 group.

We conclude this section with a statement of the representation of the solution operators for the Cauchy problem $u''(t) = (B^2 + c^2 I)u(t)$, $u(0) = \phi$, $u'(0) = \psi$. Let

$$S(t) = \tfrac{1}{2}[U(t) + U(-t)],$$

$$T(t) = \int_0^t S(\tau)\, d\tau,$$

where $U(t)$ is the C_0 group generated by B. Then

$$\tilde{S}(t) = S(t) + ct \int_0^t \frac{I_1(c\sqrt{t^2 - \sigma^2})}{\sqrt{t^2 - \sigma^2}}\, S(\sigma)\, d\sigma,$$

$$\tilde{T}(t) = T(t) + c \int_0^t \frac{I_1(c\sqrt{t^2 - \sigma^2})}{\sqrt{t^2 - \sigma^2}}\, \sigma T(\sigma)\, d\sigma,$$

where $I_1(x)$ is the modified Bessel function of order 1 and the equations are to be interpreted in the strong operator topology. These results were obtained in [10]. In terms of these operators the solution of the Cauchy problem is

$$u(t) = \tilde{S}(t)\phi + \tilde{T}(t)\psi.$$

3. Saturation theorems ($c = 0$). In this section, we consider the relevant saturation theorems for the solution operators in the case $c = 0$. We first note that the operators $S(t)$ and $T(t)$ have as domains the whole space X. Now consider the approximation process $S(t)f$, $f \in X$. This is obviously commutative and is a strong approximation process since

$$\|S(t)f-f\| \to 0$$

as $t \to 0$. Now let $f \in D(A)=D(B^2)$. Then

$$U(t)f=f+tBf+\int_0^t (t-\sigma)U(\sigma)B^2f\,d\sigma,$$

$$U(-t)f=f-tBf+\int_0^t (t-\sigma)U(-\sigma)B^2f\,d\sigma$$

and

$$\frac{S(t)f-f}{t^2}=\frac{1}{t^2}\int_0^t (t-\sigma)S(\sigma)B^2f\,d\sigma,$$

$$\left\|\frac{S(t)f-f}{t^2}\right\| \leqq \frac{1}{2}\sup_{0\leqq\sigma\leqq t}\|S(\sigma)B^2f\|$$

Therefore, for all $f \in D(B^2)$, $\|S(t)f-f\|=O(t^2)$ as $t \to 0$. Furthermore, for $f \in D(B^2)$

$$\frac{S(t)f-f}{t^2}-\frac{B^2}{2}f=\frac{1}{t^2}\int_0^t (t-\sigma)[S(\sigma)-I]B^2f\,d\sigma.$$

For sufficiently small t, $\|(S(t)-I)B^2f\|<\varepsilon$, and

$$\left\|\frac{S(t)f-f}{t^2}-\frac{B^2}{2}f\right\| \leqq \frac{\varepsilon}{t^2}\int_0^t (t-\sigma)\,d\sigma=\frac{\varepsilon}{2}.$$

Since ε is arbitrary

$$\left\|\frac{S(t)f-f}{t^2}-\frac{B^2}{2}f\right\| \to 0 \quad \text{as } t \to 0.$$

We shall want to use Theorem 13.4.1 of [8] and for this purpose we shall need a regularization process

$$J_nf=n^2\int_0^{1/n}\int_0^{1/n} U(\sigma+\eta)f\,d\sigma\,d\eta,$$

$n=1,2,3,\cdots$. Clearly $\{J_n\}$ is a family of bounded linear operators defined on X, $J_n[X]\subset D(B^2)$ for each n, $\|J_nf-f\|\to 0$ as $n\to\infty$ for each $f\in X$, and J_n commutes with $S(t)$ for each n and each t.

An appeal to Theorem 13.4.1 of [8] now yields the following result.

THEOREM 1. (i) *If $f\in X$ is such that $\|S(t)f-f\|=o(t^2)$ then $f\in D(B^2)$ and $B^2f=0$.* (ii) *$\|S(t)f-f\|=O(t^2)$ if and only if $f\in \widetilde{D(B^2)}^X$ (the relative completion of $D(B^2)$ in X). If X is reflexive then $\widetilde{D(B^2)}^X=D(B^2)$.*

The process $T(t)f$ is not an approximation process because $T(t)f\to 0$ as $t\to 0$. However, closely related to it is the process

$$V(t)f=t^{-1}T(t)f=\frac{1}{t}\int_0^t S(\tau)f\,d\tau$$

and $\|V(t)f-f\|\to 0$ as $t\to 0$. In fact, $\lim_{t\to 0} V(t)\psi$ gives the derivative $u'(0)$ of the solution of the Cauchy problem. $V(t)f$ is a commutative strong approximation process. If $f\in D(B^2)$, then

$$V(t)f-f=\frac{1}{t}\int_0^t [S(\tau)f-f]\,d\tau$$

$$=\frac{1}{t}\int_0^t \left\{\frac{\tau^2}{2}B^2f+\int_0^\tau (\tau-\sigma)[S(\sigma)-I]B^2f\,d\sigma\right\}d\tau$$

$$=\frac{t^2}{6}B^2f+\frac{1}{t}\int_0^t\int_0^\tau (\tau-\sigma)[S(\sigma)-I]B^2f\,d\sigma\,d\tau,$$

$$\left\|\frac{V(t)f-f}{t^2}-\frac{B^2}{6}f\right\|\leqq\frac{1}{t^3}\int_0^t\int_0^\tau (\tau-\sigma)\|[S(\sigma)-I]B^2f\|\,d\sigma\,d\tau.$$

For t sufficiently small $\|(S(t)-I)B^2f\|<\varepsilon$ and

$$\left\|\frac{V(t)f-f}{t^2}-\frac{B^2}{6}f\right\|<\frac{\varepsilon}{6}.$$

Since ε is arbitrary

$$\left\|\frac{V(t)f-f}{t^2}-\frac{B^2f}{6}\right\|\to 0$$

as $t\to 0$ for each $f\in D(B^2)$. At this point we introduce the same regularization process as in Theorem 1 and appealing to Theorem 13.4.1 of [8], we have

THEOREM 2. (i) *If $f\in X$ is such that $\|V(t)f-f\|=o(t^2)$, then $f\in D(B^2)$ and $B^2f=0$.* (ii) *$\|V(t)f-f\|=O(t^2)$ if and only if $f\in \widetilde{D(B^2)}^X$. If X is reflexive then $\widetilde{D(B^2)}^X=D(B^2)$.*

4. Saturation theorems ($c\neq 0$). In the case $c\neq 0$, we have since $x^{-1}I_1(x)=1/2+x^2/16+\cdots=1/2+O(x^2)$ as $x\to 0$,

$$\tilde{S}(t)f=S(t)f+c^2t\int_0^t \frac{I_1(c\sqrt{t^2-\sigma^2})}{c\sqrt{t^2-\sigma^2}}S(\sigma)f\,d\sigma$$

$$=S(t)f+c^2t^2\int_0^1 \frac{I_1(ct\sqrt{1-\eta^2})}{ct\sqrt{1-\eta^2}}S(t\eta)f\,d\eta$$

$$=S(t)f+\frac{c^2t^2}{2}\int_0^1 S(t\eta)f\,d\eta+O(t^4)$$

$$=S(t)f+\frac{c^2t}{2}\int_0^t S(\sigma)f\,d\sigma+O(t^4).$$

Then

$$\tilde{S}(t)f-f=S(t)f-f+\frac{c^2t}{2}\int_0^t S(\sigma)f\,d\sigma+O(t^4)$$

and $\|\tilde{S}(t)f-f\|\to 0$ as $t\to 0$ for each $f\in X$. Hence $\tilde{S}(t)f$ is a commutative strong approximation process. Furthermore, if $f\in D(B^2)$,

$$\frac{\tilde{S}(t)f-f}{t^2}-\frac{B^2+c^2}{2}f=\frac{S(t)f-f}{t^2}-\frac{B^2}{2}f+\frac{c^2}{2t}\int_0^t [S(\sigma)f-f]\,d\sigma+O(t^2)$$

and

$$\left\|\frac{\tilde{S}(t)f-f}{t^2}-\frac{B^2+c^2}{2}f\right\|\to 0$$

as $t\to 0$. This shows that $\|\tilde{S}(t)f-f\|=O(t^2)$ when $f\in D(B^2)$.

Next we show that in case $\|\tilde{S}(t)f-f\|=o(t^2)$ then $f\in D(B^2)$ and $(B^2+c^2)f=0$. In fact, suppose

$$\left\|\frac{\tilde{S}(t)f-f}{t^2}-g\right\|\to 0$$

as $t\to 0$. Then

$$\left\|\frac{S(t)f-f}{t^2}-g+\frac{c^2f}{2}+\frac{c^2}{2}\left(\frac{1}{t}\int_0^t S(\sigma)f\,d\sigma-f\right)\right\|\to 0$$

and therefore

$$\left\|\frac{S(t)f-f}{t^2}-\left(g-\frac{c^2}{2}f\right)\right\|\to 0.$$

This shows (see the proof of Theorem 13.4.1 of [8]) that

$$f\in D(B^2)\quad\text{and}\quad \frac{B^2}{2}f=g-\frac{c^2}{2}f\quad\text{or}\quad \frac{B^2+c^2}{2}f=g.$$

But if $g=0$ then $(B^2+c^2)f=0$.

Finally, if $\|\tilde{S}(t)f-f\|=O(t^2)$ then $\|S(t)f-f\|=O(t^2)$ and by Theorem 1 this implies that $f\in \widetilde{D(B^2)}^X$. We have then proved

THEOREM 3. (i) *If $f\in X$ is such that $\|\tilde{S}(t)f-f\|=o(t^2)$ then $f\in D(B^2)$ and $(B^2+c^2)f=0$.* (ii) *$\|\tilde{S}(t)f-f\|=O(t^2)$ if and only if $f\in \widetilde{D(B^2)}^X$.*

Finally, we consider a saturation theorem for the operator $\tilde{V}(t)=t^{-1}\tilde{T}(t)$. We have

$$\begin{aligned}\tilde{T}(t)f&=T(t)f+c^2\int_0^t \frac{I_1(c\sqrt{t^2-\sigma^2})}{c\sqrt{t^2-\sigma^2}}\sigma T(\sigma)f\,d\sigma\\ &=T(t)f+c^2t^2\int_0^1 \frac{I_1(ct\sqrt{1-\eta^2})}{ct\sqrt{1-\eta^2}}\eta T(t\eta)f\,d\eta\\ &=T(t)f+\frac{c^2t^2}{2}\int_0^1 \eta T(t\eta)f\,d\eta+O(t^4),\\ \tilde{V}(t)f&=V(t)f+\frac{c^2t}{2}\int_0^1 \eta T(t\eta)f\,d\eta+O(t^3).\end{aligned}$$

Therefore, $\tilde{V}(t)f-f=V(t)f-f+O(t)$ and $\|\tilde{V}(t)f-f\|\to 0$ as $t\to 0$, and $\tilde{V}(t)f$ is a commutative strong approximation process. Furthermore,

$$\begin{aligned}\tilde{T}(t)f&=T(t)f+\frac{c^2}{2}\int_0^t \sigma T(\sigma)f\,d\sigma+O(t^4)\\&=T(t)f+\frac{c^2}{2}\left[\frac{\sigma^2}{2}T(\sigma)f\Big|_0^t-\int_0^t\frac{\sigma^2}{2}S(\sigma)f\,d\sigma\right]+O(t^4)\\&=T(t)f+\frac{c^2t^2}{4}T(t)f-\frac{c^2}{4}\int_0^t\sigma^2S(\sigma)f\,d\sigma+O(t^4).\end{aligned}$$

Hence,

$$\frac{\tilde{V}(t)f-f}{t^2}=\frac{V(t)f-f}{t^2}+\frac{c^2}{4}V(t)f-\frac{c^2}{4}\frac{1}{t^3}\int_0^t\sigma^2S(\sigma)f\,d\sigma+O(t).$$

Now $(1/t^3)\int_0^t\sigma^2S(\sigma)f\,d\sigma\to f/3$ as $t\to 0$. In fact, $(1/t^3)\int_0^t\sigma^2 f\,d\sigma=f/3$ and for t sufficiently small $\|S(t)f-f\|<\varepsilon$. Therefore,

$$\left\|\frac{1}{t^3}\int_0^t\sigma^2[S(\sigma)f-f]\,d\sigma\right\|\leq\frac{\varepsilon}{t^3}\int_0^t\sigma^2\,d\sigma=\frac{\varepsilon}{3}.$$

If $f\in D(B^2)$ then $[V(t)f-f]/t^2\to(B^2/6)f$ and therefore we have shown that if $f\in D(B^2)$,

$$\left\|\frac{\tilde{V}(t)f-f}{t^2}-\frac{B^2+c^2}{6}f\right\|\to 0$$

as $t\to 0$. This shows that $\|\tilde{V}(t)f-f\|=O(t^2)$ for each $f\in D(B^2)$. Conversely, if $\|\tilde{V}(t)f-f\|=O(t^2)$ then $\|V(t)f-f\|=O(t^2)$ and $f\in\widetilde{D(B^2)}^X$. Finally, suppose

$$\left\|\frac{\tilde{V}(t)f-f}{t^2}-g\right\|\to 0$$

as $t\to 0$. Then

$$\left\|\frac{V(t)f-f}{t^2}-\left(g-\frac{c^2}{6}f\right)\right\|\to 0$$

as $t\to 0$, which implies that $f\in D(B^2)$ and $(B^2/6)f=g-(c^2/6)f$ or $[(B^2+c^2)/6]f=g$. If $g=0$ then $(B^2+c^2)f=0$. We have proved

THEOREM 4. (i) *If $f\in X$ is such that $\|\tilde{V}(t)f-f\|=o(t^2)$ then $f\in D(B^2)$ and $(B^2+c^2)f=0$.* (ii) *$\|\tilde{V}(t)f-f\|=O(t^2)$ if and only if $f\in\widetilde{D(B^2)}^X$.*

5. Orders of approximation in the Cauchy problem. If $\phi\in D(B^2)$ and $\psi\in D(B)\cap R(B)$, then we can write the solution of the Cauchy problem as

$$u(t)=\tilde{S}(t)\phi+\tilde{T}(t)\psi.$$

In fact, this expression makes perfectly good sense for any $\phi,\psi\in X$. Therefore, we shall refer to it as the generalized solution of the Cauchy problem. We now consider the order of magnitude of $\|u(t)-\phi\|$ as $t\to 0$.

THEOREM 5. *If $\phi \in X$ is such that $\|u(t)-\phi\| = o(t^2)$, then $\psi = 0$, $\phi \in D(B^2)$, $(B^2+c^2)\phi = 0$, and $u(t) \equiv \phi$. Furthermore, $\|u(t)-\phi\| = O(t^2)$ if and only if $\psi = 0$ and $\phi \in \widetilde{D(B^2)}^X$.*

Proof. If $\psi = 0$ then $u(t)-\phi = \tilde{S}(t)\phi - \phi$ and $\|\tilde{S}(t)\phi - \phi\| = \|u(t)-\phi\| = O(t^2)$ when $\phi \in \widetilde{D(B^2)}^X$. Conversely, suppose that $\|u(t)\phi - \phi\| = O(t^2)$; then

$$0 = \lim_{t\to 0} \frac{u(t)-\phi}{t} = \lim_{t\to 0} \frac{S(t)\phi - \phi}{t} + \lim_{t\to 0} \tilde{V}(t)\psi.$$

Since, $\lim_{t\to 0} \tilde{V}(t)\psi = \psi$ we know that $\lim_{t\to 0} (S(t)\phi - \phi)/t$ exists. We shall show that this limit is zero and, therefore, that $\psi = 0$. For this purpose we introduce the regularization process J_n defined by

$$J_n f = n \int_0^{1/n} U(t) f \, dt.$$

We know that $\lim_{t\to 0} (S(t)f - f)/t = 0$ if $f \in D(B)$, and therefore $\lim_{t\to 0} (S(t)J_n\phi - J_n\phi/t = 0$. But J_n commutes with $S(t)$ for each n and each t, and therefore if $g = lim_{t\to 0} (S(t)\phi - \phi)/t$ then $J_n g = 0$. However, $\lim_{n\to\infty} J_n g = g = 0$. We have shown that $\psi = 0$ and hence $\|u(t)-\phi\| = \|\tilde{S}(t)\phi - \phi\| = O(t^2)$. By Theorem 3, $\phi \in \widetilde{D(B^2)}^X$. If $\|u(t)-\phi\| = o(t^2)$, $\phi \in D(B^2)$ and $(B^2+c^2)\phi = 0$. In this case, $u''(t) = (B^2+c^2)u(t) = (B^2+c^2)\tilde{S}(t)\phi = \tilde{S}(t)(B^2+c^2)\phi = 0$. Therefore, since $u'(0) = \psi = 0$, $u(t) \equiv \phi$. This completes the proof.

In order to differentiate $u(t)$ we must have that $\phi \in D(B)$, and in this case

$$u'(t) = \frac{1}{2}U(t)B\phi - \frac{1}{2}U(-t)B\phi + \frac{c^2}{2}\int_0^t S(\tau)\phi \, d\tau + \frac{c^2 t}{2} S(t)\phi + \tilde{S}(t)\psi + O(t^3).$$

If in addition, $\phi \in D(B^2)$,

$$\frac{1}{2}U(t)B\phi - \frac{1}{2}U(-t)B\phi = \int_0^t S(\tau)B^2\phi \, d\tau$$

and

$$\begin{aligned} \frac{c^2 t}{2} S(t)\phi &= \frac{c^2 t}{2}\phi + O(t^3) \\ &= \frac{c^2}{2}\int_0^t S(\tau)\phi \, d\tau + O(t^3). \end{aligned}$$

Therefore, if $\phi \in D(B^2)$ and $(B^2+c^2)\phi = 0$,

$$u'(t) - \psi = \tilde{S}(t)\psi - \psi + O(t^3).$$

Then if $\psi \in \widetilde{D(B^2)}^X$, $\|u'(t)-\psi\| = O(t^2)$. On the other hand, if $\|u'(t)-\psi\| = O(t^2)$ then

$$0 = \lim_{t\to 0} \frac{u'(t)-\psi}{t} = \lim_{t\to 0} \left[\frac{U(t)B\phi - U(-t)B\phi}{2t} + \frac{S(t)\psi - \psi}{t} + c^2\phi\right]$$

The regularization process J_n introduced in the proof of Theorem 5 commutes with $U(t)$, $U(-t)$, and $S(t)$. Therefore,

$$\lim_{t\to 0}\left[\frac{U(t)J_nB\phi - U(-t)J_nB\phi}{2t}+\frac{S(t)J_n\psi - J_n\psi}{t}\right] = -c^2 J_n\phi.$$

But $J_n\phi \in D(B)$ and $J_n\psi \in D(B)$ and hence

$$\lim_{t\to 0}\frac{U(t)J_nB\phi - U(-t)J_nB\phi}{2t} = BJ_nB\phi,$$

$$\lim_{t\to 0}\frac{S(t)J_n\psi - J_n\psi}{t} = 0.$$

The operator B is closed, which implies that $\phi \in D(B^2)$ and

$$\lim_{n\to\infty}(BJ_nB\phi + c^2 J_n\phi) = (B^2+c^2)\phi = 0.$$

From this we have $u'(t)-\psi = \tilde{S}(t)\psi - \psi + O(t^3)$ and Theorem 3 implies that $\psi \in \widetilde{D(B^2)}^X$. If $\|u'(t)-\psi\| = o(t^2)$ then $\psi \in D(B^2)$ and $(B^2+c^2)\psi = 0$. In this case,

$$u''(t) = (B^2+c^2)u(t) = \tilde{S}(t)(B^2+c^2)\phi + \tilde{T}(t)(B^2+c^2)\psi = 0$$

and $u(t) = \phi + t\psi$. We have proved

THEOREM 6. *If $\phi \in D(B)$ and ψ are such that $\|u'(t)-\psi\| = o(t^2)$ then $\phi \in D(B^2)$, $\psi \in D(B^2)$, $(B^2+c^2)\phi = (B^2+c^2)\psi = 0$, and $u(t) = \phi + t\psi$. Furthermore, $\|u'(t)-\psi\| = O(t^2)$ if and only if $\phi \in D(B^2)$, $(B^2+c^2)\phi = 0$, and $\psi \in \widetilde{D(B^2)}^X$.*

In order not to have to assume at the outset that $\phi \in D(B)$, it is of interest to consider the difference quotient $t^{-1}[u(t)-\phi]$ as an approximation process for $u'(0) = \psi$. We therefore consider

$$\frac{u(t)-\phi}{t} - \psi = \frac{\tilde{S}(t)\phi - \phi}{t} + \tilde{V}(t)\psi - \psi.$$

If $\phi \in D(B^2)$ and $(B^2+c^2)\phi = 0$, then $\tilde{u}(t) = \tilde{S}(t)\phi$ is the unique solution of $\tilde{u}''(t) = A\tilde{u}(t)$, $\tilde{u}(0) = \phi$, $\tilde{u}'(0) = 0$. Then $\tilde{u}''(t) = (B^2+c^2)\tilde{S}(t)\phi = \tilde{S}(t)(B^2+c^2)\phi = 0$. Hence $\tilde{u}(t) = \tilde{S}(t)\phi \equiv \phi$. This shows that

$$\frac{u(t)-\phi}{t} - \psi = \tilde{V}(t)\psi - \psi.$$

Therefore, if $\psi \in \widetilde{D(B^2)}^X$, $\|\tilde{V}t)\psi - \psi\| = O(t^2)$ and

$$\left\|\frac{u(t)-\phi}{t} - \psi\right\| = O(t^2).$$

Conversely, suppose $\|t^{-1}[u(t)-\phi] - \psi\| = O(t^2)$; then

$$0 = \lim_{t\to 0}\left[\frac{u(t)-\phi}{t^2} - \frac{\psi}{t}\right] = \lim_{t\to 0}\left[\frac{\tilde{S}(t)\phi - \phi}{t^2} + \frac{\tilde{V}(t)\psi - \psi}{t}\right]$$

and

$$\lim_{t\to 0}\left[\frac{S(t)\phi - \phi}{t^2} + \frac{V(t)\psi - \psi}{t}\right] = -\frac{c^2}{2}\phi.$$

At this point we introduce the regularization process

$$J_nf = n^2 \int_0^{1/n} \int_0^{1/n} U(\sigma+\eta)f\, d\sigma\, d\eta$$

used in the proof of Theorem 1. J_n commutes with both $S(t)$ and $V(t)$ and $J_nf \in D(B^2)$ for all $f \in X$. Therefore,

$$\lim_{t\to 0}\left[\frac{S(t)J_n\phi - J_n\phi}{t^2} + \frac{V(t)J_n\psi - J_n\psi}{t}\right] = -\frac{c^2}{2}J_n\phi.$$

Since $J_n\psi \in D(B^2)$, $\lim_{t\to 0}(V(t)J_n\psi - J_n\psi)/t = 0$ and this implies that

$$\lim_{t\to 0}\frac{(S(t)J_n\phi - J_n\phi)}{t^2} = \frac{B^2}{2}J_n\phi = -\frac{c^2}{2}J_n\phi$$

or $(B^2+c^2)J_n\phi = 0$. The operator B^2 is closed and hence $\phi \in D(B^2)$ and $(B^2+c^2)\phi = 0$. From this it follows that $\|\tilde{V}(t)\psi - \psi\| = O(t^2)$ and $\psi \in \widetilde{D(B^2)}^X$. If $\|t^{-1}[u(t)-\phi]-\psi\| = o(t^2)$ then $\|\tilde{V}(t)\psi - \psi\| = o(t^2)$ and by Theorem 4, $\psi \in D(B^2)$ and $(B^2+c^2)\psi = 0$. Then $u''(t) = (B^2+c^2)u(t) = \tilde{S}(t)(B^2+c^2)\phi + \tilde{T}(t)(B^2+c^2)\psi = 0$ and $u(t) = \phi + t\psi$. We have proved

THEOREM 7. *If ϕ and ψ are such that $\|t^{-1}[u(t)-\phi]-\psi\| = o(t^2)$ then $\phi \in D(B^2)$, $\psi \in D(B^2)$, $(B^2+c^2)\phi = (B^2+c^2)\psi = 0$ and $u(t) = \phi + t\psi$. Furthermore, $\|t^{-1}[u(t)-\phi]-\psi\| = O(t^2)$ if and only if $\phi \in D(B^2)$, $(B^2+c^2)\phi = 0$, and $\psi \in \widetilde{D(B^2)}^X$.*

REFERENCES

[1] H. ARONSZAJN AND E. GAGLIARDO, *Interpolation spaces and interpolation methods*, Ann. Mat. Pura. Appl. (4), 68 (1965), pp. 51–118.

[2] A. V. BALAKRISHNAN, *Fractional powers of closed operators and the semigroups generated by them*, Pacific J. Math., 10 (1960), pp. 419–437.

[3] H. BERENS, *Interpolation methoden zur Behandlung von Approximationsprozessen auf Banachräumen*, Lecture Notes in Mathematics 64, Springer-Verlag, Berlin, 1968.

[4] L. R. BRAGG AND J. W. DETTMAN, *Related problems in partial differential equations*, Bull. Amer. Math. Soc., 74 (1967), pp. 428–430.

[5] ———, *An operator calculus for related partial differential equations*, J. Math. Anal. Appl., 22 (1968), pp. 261–271.

[6] ———, *Related partial differential equations and their applications*, SIAM J. Appl. Math., 16 (1968), pp. 459–467.

[7] P. L. BUTZER AND H. BERENS, *Semi-groups of Operators and Approximation*, Springer-Verlag, New York, 1967.

[8] P. L. BUTZER AND R. J. NESSEL, *Fourier Analysis and Approximations*, vol. I, Academic Press, New York, 1971.

[9] G. DAPRATO AND E. GIUSTI, *Una caratterizzazione dei generatori di funzioni cosen astratte*, Boll. Un Mat. Ital., 22 (1967), pp. 357–362.

[10] J. W. DETTMAN, *Perturbation techniques in related differential equations*, J. Differential Equations, 14 (1973), pp. 547–558.

[11] H. O. FATTORINI, *Ordinary differential equations in linear topological spaces. I*, Ibid., 5 (1968), pp. 72–105.

[12] ———, *Ordinary differential equations in linear topological spaces II*, Ibid., 6 (1969), pp. 50–70.

[13] E. GAGLIARDO, *A unified structure in various families of function spaces. Compactness and closure theorems*, Symposium on Linear Spaces (Hebrew Univ., Jerusalem, 1960), Jerusalem Academic Press–Pergamon Press, Jerusalem, 1961, pp. 237–241.

[14] J. A. GOLDSTEIN, *On a connection between first and second order differential equations in Banach spaces*, J. Math. Anal. Appl., 30 (1970), pp. 246–251.

[15] E. HILLE AND R. S. PHILLIPS, *Functional Analysis and Semi-groups*, Amer. Math. Soc. Colloquium Publications, vol. XXXI, Providence, R.I., 1957.

[16] H. S. SHAPIRO, *Smoothing and Approximation of Functions*, Van Nostrand, New York, 1969.

[17] M. SOVA, *Cosine operator functions*, Rozprawy Mat., 49 (1966), pp. 1–47.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

SOME COMBINATORIAL IDENTITIES OF BERNSTEIN*

L. CARLITZ†

Abstract. For g a rational integer such that $\Delta = 4g^3 + 27$ is square-free, let w denote the real root of $u^3 + gu - 1 = 0$ and put $w^n = r_n + s_n w + t_n w^2$, $w^{-n} = x_n + y_n w + z_n w^2$, $n \geqq 0$. Making use of the theory of units in an algebraic number field, Bernstein obtained quadratic relations involving the r_n and x_n as well as explicit formulas. These lead to certain combinatorial identities. In the present paper these and related identities are proved using only some elementary algebra.

1. Introduction. Let g be a rational integer such that $\Delta = 4g^3 + 27$ is square-free and let w denote the real root of the irreducible equation

$$x^3 + gx - 1 = 0 \qquad (g > 1). \tag{1.1}$$

Clearly w is a unit of the cubic field $Q(w)$. Put

$$w^n = r_n + s_n w + t_n w^2 \qquad (n \geqq 0) \tag{1.2}$$

and

$$w^{-n} = x_n + y_n w + z_n w^2 \qquad (n \geqq 0). \tag{1.3}$$

Making use of the theory of units in an algebraic number field, Bernstein [2], [3] has obtained certain combinatorial identities. He showed that

$$\sum_{n=0}^{\infty} r_n u^n = \frac{1 + gu^2}{1 + gu^2 - u^3} \tag{1.4}$$

and

$$\sum_{n=0}^{\infty} x_n u^n = \frac{1}{1 - gu - u^3}. \tag{1.5}$$

It follows from (1.4) and (1.5) that

$$\begin{aligned} r_{2n} &= \sum_{3k \leqq n} (-1)^{n-k} \binom{n-k-1}{2k-1} g^{n-3k}, \\ r_{2n+1} &= \sum_{3k < n} (-1)^{n-k-1} \binom{n-k-1}{2k} g^{n-3k-1} \end{aligned} \tag{1.6}$$

and

$$\begin{aligned} x_{3n} &= \sum_{k=0}^{n} \binom{n+2k}{3k} g^{3k}, \\ x_{3n+1} &= \sum_{k=0}^{n} \binom{n+2k+1}{3k+1} g^{3k+1}, \\ x_{3n+2} &= \sum_{k=0}^{n} \binom{n+2k+2}{3k+2} g^{3k+2}. \end{aligned} \tag{1.7}$$

---

* Received by the editors April 30, 1976, and in revised form August 10, 1976.

† Department of Mathematics, Duke University, Durham, North Carolina 27706. This work was supported in part by the National Science Foundation under Grant GP-37924X.

Moreover, by (1.2) and (1.3),

$$r_n^2 - r_{n-1}r_{n+1} = x_{n-3}, \qquad (1.8)$$
$$x_n^2 - x_{n-1}x_{n+1} = r_{n+3}.$$

Substituting from (1.6) and (1.7) in (1.8), we obtain the combinatorial identities. Since $\Delta = 4g^3 + 27$ is squarefree for infinitely many values of g, the identities are indeed polynomial identities.

The object of the present paper is to prove these and related identities using only some elementary algebra. Let g denote an indeterminate and put

$$1 + gx^2 - x^3 = (1-\alpha x)(1-\beta x)(1-\gamma x).$$

Put

$$\sigma_n = \alpha^n + \beta^n + \gamma^n, \qquad (1.9)$$

where n is an arbitrary positive or negative integer. Then

$$3r_n + gr_{n-2} = \sigma_n + g\sigma_{n-2} \qquad (n \geqq 2)$$

and

$$\sigma_{-n} = 3x_n - 2gx_{n-1} \qquad (n \geqq 1).$$

We shall show that

$$\sigma_{2n} = \sigma_n^2 - \sigma_{-n}$$

and that generally σ_{kn}, $k \geqq 1$, is a polynomial in σ_n, σ_{-n} with integral coefficients; see (3.14) for an explicit result. Moreover

$$\sigma_m\sigma_n = \sigma_{m+n} + \sigma_{m-n}\sigma_{-n} - \sigma_{m-2n}, \qquad (1.10)$$

for arbitrary m, n. There are numerous formulas involving the products $r_m\sigma_n$, $x_m\sigma_n$. If we let $\rho_n = r_n$ or x_{-n} according as $n \geqq 0$ or $n \leqq 0$, these formulas can be included in a single identity:

$$\rho_m\sigma_n = \rho_{m+n} + \rho_{m-n}\sigma_{-n} - \rho_{m-2n} \qquad (1.11)$$

and

$$2\rho_m\rho_n - \rho_{m+1}\rho_{n-1} - \rho_{m-1}\rho_{n+1} \qquad (1.12)$$
$$= \sigma_{m-3}\sigma_{n-3} - \sigma_{m+3-6} - \sigma_{m-3}\rho_{n-3} - \sigma_{n-3}\rho_{m-3} + 2\rho_{m+n-6}$$

for arbitrary m,n. The latter formula contains (1.8) as a special case.

We also obtain an explicit formula for ρ_{kn} in terms of σ_n, σ_{-n}, ρ_n, ρ_{-n}.

Corresponding to (1.6) and (1.7) we have

$$\sigma_n = \sum_{n/3 \leqq k \leqq n/2} (-1)^{n-k} \frac{n}{k} \binom{n}{n-2k} g^{3k-n} \qquad (n>0) \qquad (1.13)$$

and

$$\sigma_{-n} = \sum_{3j \leqq n} \frac{n}{n-2j} \binom{n-2j}{j} g^{n-3j} \qquad (n>0). \qquad (1.14)$$

The functions r_n, x_n, σ_n, σ_{-n} are all polynomials in the indeterminate g and may be thought of as analogues of the Chebyshev polynomials [4, Chap. 4]. For a detailed discussion of the linearization of the product of two classical orthogonal polynomial see Askey [1, Lecture 5].

2. In what follows g will denote an indeterminate. Put

$$1+gx^2-x^3=(1-\alpha x)(1-\beta x)(1-\gamma x), \tag{2.1}$$

so that α, β, γ are the roots of

$$z^3+gz-1=0. \tag{2.2}$$

Let

$$\frac{1+gx^2}{1+gx^2-x^3}=\frac{A}{1-\alpha x}+\frac{B}{1-\beta x}+\frac{C}{1-\gamma x}, \tag{2.3}$$

where A, B, C are independent of x. By (2.1)

$$\begin{gathered}\alpha+\beta+\gamma=0,\\ \beta\gamma+\gamma\alpha+\alpha\beta=g, \qquad \alpha\beta\gamma=1\end{gathered} \tag{2.4}$$

and by (2.3)

$$A=\frac{g+\alpha^2}{(\alpha-\beta)(\alpha-\gamma)}.$$

It follows from (2.4) that

$$A=\frac{1}{3-2g\alpha} \tag{2.5}$$

with similar formulas for B and C.

Comparing (1.4) with (2.3) we get

$$r_n=\sum A\alpha^n\equiv A\alpha^n+B\beta^n+C\gamma^n \qquad (n\geqq 0). \tag{2.6}$$

It is easily verified that

$$\Delta=4g^3+27=\prod(3-2g\alpha)=\prod A^{-1}. \tag{2.7}$$

Hence, by (2.6),

$$\begin{aligned}\Delta\cdot r_n&=\sum(3-2g\beta)(3-2g\gamma)\alpha^n\\ &=\sum(9+6g\alpha+4g^2\beta\gamma)\alpha^n.\end{aligned}$$

We define

$$\sigma_n=\sum\alpha^n\equiv\alpha^n+\beta^n+\gamma^n \qquad (n\in Z). \tag{2.8}$$

Thus

$$\Delta\cdot r_n=9\sigma_n+6g\sigma_{n+1}+4g^2\sigma_{n-1}. \tag{2.9}$$

In the next place, since

$$1-gu-u^3=(1-\beta\gamma u)(1-\gamma\alpha u)(1-\alpha\beta u),$$

we may put

$$\frac{1}{1-gu-u^3}=\frac{A'}{1-\beta\gamma u}+\frac{B'}{1-\gamma\alpha u}+\frac{C'}{1-\alpha\beta u}, \tag{2.10}$$

where A', B', C' are independent of u. Then

$$1=\sum A'(1-\gamma\alpha u)(1-\alpha\beta u).$$

For $u=\alpha$ this reduces to

$$\begin{aligned}1&=A'(1-\alpha^2\beta)(1-\alpha^2\gamma)\\&=A'(1-\alpha^2(\beta+\gamma)+\alpha^4\beta\gamma)\\&=A'(1+2\alpha^3)=A'(3-2g\alpha),\end{aligned}$$

so that

$$A'=A,\quad B'=B,\quad C'=C. \tag{2.11}$$

Hence (2.10) becomes

$$\frac{1}{1-gu-u^3}=\sum\frac{A}{1-\beta\gamma u}. \tag{2.12}$$

Comparison with (1.5) gives

$$x_n=\sum A\beta^n\gamma^n=\sum A\alpha^{-n}\qquad(n\geqq 0). \tag{2.13}$$

It follows from (2.6) that

$$r_m r_n=\sum A^2\alpha^{m+n}+\sum BC(\beta^m\gamma^n+\gamma^m\beta^n). \tag{2.14}$$

In particular

$$\begin{aligned}r_n^2&=\sum A^2\alpha^{2n}+2\sum BC\beta^n\gamma^n,\\r_{n+1}r_{n-1}&=\sum A^2\alpha^{2n}+\sum BC\beta^{n-1}\gamma^{n-1}(\beta^2+\gamma^2).\end{aligned}$$

Subtracting the second equation from the first, we get

$$r_n^2-r_{n+1}r_{n-1}=-\sum BC\beta^{n-1}\gamma^{n-1}(\beta-\gamma)^2.$$

It is easily verified that

$$BC(\beta-\gamma)^2=-A\alpha^2,$$

so that

$$r_n^2-r_{n+1}r_{n-1}=\sum A\beta^{n-3}\gamma^{n-3}. \tag{2.15}$$

Therefore, by (2.13),

(2.16)

$$r_n^2-r_{n+1}r_{n-1}=x_{n-3},$$

in agreement with the first of (1.8).

Next, by (2.13),

$$x_m x_n = \sum A^2\beta^{m+n}\gamma^{m+n} + \sum BC\alpha^{m+n}(\beta^m\gamma^n + \gamma^m\beta^n). \tag{2.17}$$

Hence, exactly as above, we get

$$\begin{aligned} x_n^2 - x_{n+1}x_{n-1} &= -\sum BC\alpha^{2n}\beta^{n-1}\gamma^{n-1}(\beta-\gamma)^2 \\ &= \sum A\alpha^{2n+2}\beta^{n-1}\gamma^{n-1} = \sum A\alpha^{n+3}, \end{aligned}$$

so that

$$x_n^2 - x_{n+1}x_{n-1} = r_{n+3}, \tag{2.18}$$

in agreement with the second of (1.8).

The coefficients A^2, BC in (2.14) suggest that it may be of interest to evaluate

$$\sum \frac{A^2}{1-\alpha x},\quad \sum \frac{BC}{1-\alpha x},\quad \sum \frac{A^2}{1-\beta\gamma x},\quad \sum \frac{BC}{1-\beta\gamma x}.$$

We find that

$$\sum \frac{A^2}{1-\alpha x} = \frac{1}{\Delta^2}\frac{a+bx+cx^2}{1+gx^2-x^3}, \tag{2.19}$$

where

$$\begin{aligned} a &= 243+144g^2+16g^6, \\ b &= 4(108-81g^2+4g^5), \\ c &= 3g(135-48g+24g^3+16g^6); \end{aligned}$$

$$\sum \frac{BC}{1-\alpha x} = \frac{1}{\Delta}\frac{27+4g^2x-3gx^2}{1+gx^2-x^3}; \tag{2.20}$$

$$\sum \frac{A^2}{1-\beta\gamma x} = \frac{1}{\Delta^2}\frac{a+6g(27+16g^2)x+bx^2}{1-gx-x^3} \tag{2.21}$$

where a, b have the same meaning as in (2.19);

$$\sum \frac{BC}{1-\beta\gamma x} = \frac{1}{\Delta}\frac{27-3gx+4g^2x^2}{1-gx-x^3}. \tag{2.22}$$

3. We have defined

$$\sigma_n = \sum \alpha^n, \tag{3.1}$$

for all integral n. Thus

$$\sum_{n=0}^{\infty} \sigma_n x^n = \sum \frac{1}{1-\alpha x} = \sum \frac{(1-\beta x)(1-\gamma x)}{1+gx^2-x^3},$$

which reduces to

$$\sum_{n=0}^{\infty} \sigma_n x^n = \frac{3+gx^2}{1+gx^2-x^3}. \tag{3.2}$$

Similarly we have

(3.3)
$$\sum_{n=0}^{\infty} \sigma_{-n}x^n = \frac{3-2gx}{1-gx-x^3}.$$

Comparing (3.2) with (1.4) and (3.3) with (1.5), we get

(3.4)
$$3r_n + gr_{n-2} = \sigma_n + g\sigma_{n-2} \qquad (n \geqq 2)$$

and

(3.5)
$$\sigma_{-n} = 3x_n - 2gx_{n-1} \qquad (n \geqq 1).$$

In the next place, it follows at once from (3.1) that

$$\sigma_n^2 = \sigma_{2n} + 2\sum \beta^n \nu^n.$$

Since $\alpha\beta\gamma = 1$, this gives

(3.6)
$$\sigma_n^2 = \sigma_{2n} + 2\sigma_{-n},$$

for both positive and negative n. Then

$$\tfrac{1}{4}(\sigma_n^2 - \sigma_{2n})^2 - \tfrac{1}{4}(\sigma_{2n}^2 - \sigma_{4n}) = \sigma_{-n}^2 - \sigma_{-2n} = 2\sigma_n,$$

so that

(3.7)
$$(\sigma_n^2 - \sigma_{2n})^2 = 2(\sigma_{2n}^2 - \sigma_{4n}) + 8\sigma_n.$$

We have, for arbitrary m and n,

$$\sigma_m\sigma_n = \sum \alpha^{m+n} + \sum \beta^n\gamma^n(\beta^{m-n} + \gamma^{m-n}).$$

Since

$$\sum \beta^n\gamma^n(\beta^{m-n} + \gamma^{m-n}) = \sum \beta^n\gamma^n(\sigma_{m-n} - \alpha^{m-n})$$
$$= \sigma_{m-n}\sigma_{-n} - \sum (\alpha\beta\gamma)^n\alpha^{m-2n},$$

it follows that

(3.8)
$$\sigma_m\sigma_n = \sigma_{m+n} + \sigma_{m-n}\sigma_{-n} - \sigma_{m-2n}.$$

For $m = n$, this reduces to (3.6).

Interchanging m, n in (3.8) and subtracting the result from (3.8), we get

(3.9)
$$\sigma_{m-n}\sigma_{-n} - \sigma_{m-2n} = \sigma_{n-m}\sigma_{-m} - \sigma_{n-2m}.$$

A slightly more symmetrical form is obtained on replacing n by $-n$:

(3.10)
$$\sigma_{m+n}\sigma_n - \sigma_{m+2n} = \sigma_{-m-n}\sigma_{-m} - \sigma_{-2m-n}.$$

For $m = 2n$, (3.8) reduces to

$$\sigma_{2n}\sigma_n = \sigma_{3n} + \sigma_n\sigma_{-n} - 3.$$

Thus by (3.6)

(3.11)
$$\sigma_{3n} = \sigma_n^3 - 3\sigma_n\sigma_{-n} + 3.$$

For $m = 3n$, we get

$$\sigma_{4n} = \sigma_{3n}\sigma_n - \sigma_{2n}\sigma_{-n} + \sigma_n,$$

so that

$$\sigma_{4n} = \sigma_n^4 - 4\sigma_n^2\sigma_{-n} + 4\sigma_n + 2\sigma_{-n}^2. \tag{3.12}$$

Similarly

$$\sigma_{5n} = \sigma_n^5 - 5\sigma_n^3\sigma_{-n} + 5\sigma_n^2 + 5\sigma_n\sigma_{-n}^2 - 5\sigma_{-n}. \tag{3.13}$$

By an easy induction, σ_{kn} is a polynomial in σ_n, σ_{-n} of degree k in σ_n and of degree $[k/2]$ in σ_{-n} and with leading term σ_n^k. Moreover one might guess that, for k prime, all coefficients except the first are divisible by k.

Clearly, for arbitrary n, we have by (2.8)

$$\sum_{k=0}^{\infty} \sigma_{kn}x^k = \sum \frac{1}{1-\alpha^n x}$$
$$= \frac{\sum (1-\beta^n x)(1-\gamma^n x)}{(1-\alpha^n x)(1-\beta^n x)(1-\gamma^n x)} = \frac{3 - 2\sigma_n x + \sigma_{-n}x^2}{1 - \sigma_n x + \sigma_{-n}x^2 - x^3}.$$

From this it is clear that σ_{kn} is a polynomial in σ_n, σ_{-n} with integral coefficients. Since

$$(1 - \sigma_n x + \sigma_{-n}x^2 - x^3)^{-1} = \sum_{r=0}^{\infty} (\sigma_n x - \sigma_{-n}x^2 + x^3)^r$$
$$= \sum_{r,s,t=0}^{\infty} (-1)^2 \frac{(r+s+t)!}{r!s!t!} \sigma_n^r \sigma_{-n}^s x^{r+2s+3t}$$
$$= \sum_{k=0}^{\infty} x^k \sum_{r+2s+3t=k} (-1)^s \frac{(r+s+t)!}{r!s!t!} \sigma_n^r \sigma_{-n}^s$$
$$= \sum_{k=0}^{\infty} x^k c_{n,k},$$

say. Thus

$$\sigma_{kn} = 3c_{n,k} - 2\sigma_n c_{n,k-1} + \sigma_{-n}c_{n,k-2}.$$

After some manipulation we find that

$$\sigma_{kn} = \sum_{r+2s+3t=k} (-1)^s \frac{k}{r+s+t} \frac{(r+s+t)!}{r!s!t!} \sigma_n^r \sigma_{-n}^s. \tag{3.14}$$

We have accordingly found an explicit formula for σ_{kn} as a polynomial in σ_n, σ_{-n}. Moreover, for k prime, it is clear that all coefficients except for the term σ_n^k are indeed divisible by k.

4. It is convenient to define

$$\rho_n = \sum A\alpha^n, \tag{4.1}$$

for all integral n. Thus

$$\rho_n = \begin{cases} r_n & (n \geqq 0), \\ x_{-n} & (n \geqq 0). \end{cases} \tag{4.2}$$

By (3.1) and (4.1), for arbitrary m and n,

$$\begin{aligned} \rho_m\sigma_n &= \sum A\alpha^m \cdot \sum \alpha^n \\ &= \sum A\alpha^{m+n} + \sum A\alpha^m(\beta^m + \gamma^n) \\ &= \rho_{m+n} + \sum A\alpha^{m-n}(\gamma^{-n} + \beta^{-n}) \\ &= \rho_{m+n} + \sum A\alpha^{m-n}(\sigma_{-n} - \alpha^{-n}). \end{aligned}$$

Hence we have, for all m and n,

$$\rho_m\sigma_n = \rho_{m+n} + \rho_{m-n}\sigma_{-n} - \rho_{m-2n}. \tag{4.3}$$

In terms of r_k and x_k, (4.3) includes numerous formulas. In particular, we have

$$r_m\sigma_n = r_{m+n} + r_{m-n}\sigma_{-n} - r_{m-2n} \qquad (m \geqq 2n \geqq 0) \tag{4.4}$$

$$x_m\sigma_{-n} = x_{m+n} + x_{m-n}\sigma_n - x_{m-2n} \qquad (m \geqq 2n \geqq 0). \tag{4.5}$$

We have also, using (4.1),

$$\begin{aligned} \sum_{k=0}^{\infty} \rho_{kn}x^k &= \sum_{k=0}^{\infty} x^k \sum A\alpha^{kn} \\ &= \sum \frac{A}{1-\alpha^n x} = \frac{\sum A(1-\beta^n x)(1-\gamma^n x)}{(1-\alpha^n x)(1-\beta^n x)(1-\gamma^n x)}. \end{aligned}$$

Since

$$(1-\alpha^n x)(1-\beta^n x)(1-\gamma^n x) = 1 - \sigma_n x + \sigma_{-n}x^2 - x^3$$

and

$$\begin{aligned} \sum A(1-\beta^n x)(1-\gamma^n x) &= \sum A - x\sum A(\sigma_n - \alpha^n) + x^2 \sum A\alpha^{-n} \\ &= 1 - (\sigma_n - \rho_n)x + \rho_{-n}x^2 = 1 + \sigma_{-n}x^2, \end{aligned}$$

we have

$$\sum_{k=0}^{\infty} \rho_{kn}x^k = \frac{1-(\sigma_n - \rho_n)x + \rho_{-n}x^2}{1 - \sigma_n x + \sigma_{-n}x^2 - x^3}. \tag{4.6}$$

Hence, as in the proof of (3.14),

$$\rho_{kn} = c_{n,k} - (\sigma_n - \rho_n)c_{n,k-1} + \rho_{-n}c_{n,k-2}, \tag{4.7}$$

where

$$c_{n,k} = \sum_{r+2s+3t=k} (-1)^s \frac{(r+s+t)!}{r!s!t!} \sigma_n^r \sigma_{-n}^s. \tag{4.8}$$

Thus we have an explicit formula for ρ_{kn} in terms of σ_n, σ_{-n}, ρ_n, ρ_{-n}.

We shall now obtain a formula for

$$R_{m,n} \equiv 2\rho_m\rho_n - \rho_{m+1}\rho_{n-1} - \rho_{m-1}\rho_{n+1}. \tag{4.9}$$

We have

$$\rho_m\rho_n = \sum A^2\alpha^{m+n} + \sum BC(\beta^m\gamma^n + \gamma^m\beta^n).$$

This gives

$$\rho_m\rho_n - \rho_{m+1}\rho_{n-1} = -\sum BC(\beta-\gamma)(\beta^m\gamma^{n-1} - \gamma^m\beta^{n-1}),$$
$$\rho_m\rho_n - \rho_{m-1}\rho_{n+1} = -\sum BC(\beta-\gamma)(\beta^n\gamma^{m-1} - \gamma^n\beta^{m-1}).$$

Also as above

$$BC(\beta-\gamma)^2 = -A\alpha^2.$$

Hence

$$\begin{aligned} R_{m,n} &= -\sum BC(\beta-\gamma)^2(\beta^{m-1}\gamma^{n-1} + \gamma^{m-1}\beta^{n-1}) \\ &= \sum A(\beta^{m-3}\gamma^{n-3} + \gamma^{m-3}\beta^{n-3}). \end{aligned} \tag{4.10}$$

Since

$$\begin{aligned} \beta^m\gamma^n + \beta^n\gamma^m &= (\beta^m+\gamma^m)(\beta^n+\gamma^n) - (\beta^{m+n}+\gamma^{m+n}) \\ &= (\sigma_m - \alpha^m)(\sigma_n - \alpha^n) - (\sigma_{m+n} - \alpha^{m+n}) \\ &= \sigma_m\sigma_n - \sigma_m\alpha^n - \sigma_n\alpha^m - \sigma_{m+n} + 2\alpha^{m+n}, \end{aligned}$$

it follows that

$$\sum A(\beta^m\gamma^n + \beta^n\gamma^m) = \sigma_m\sigma_n - \sigma_{m+n} - \sigma_m\rho_n - \sigma_n\rho_m + 2\rho_{m+n}. \tag{4.11}$$

Hence, by (4.9), (4.10) and (4.11), we have

$$\begin{aligned} &2\rho_m\rho_n - \rho_{m+1}\rho_{n-1} - \rho_{m-1}\rho_{n+1} \\ &\qquad = \sigma_{m-3}\sigma_{n-3} - \sigma_{m+n-6} - \sigma_{m-3}\rho_{n-3} - \sigma_{n-3}\rho_{m-3} + 2\rho_{m+n-6}. \end{aligned} \tag{4.12}$$

It can be verified that, for $m=n$, (4.12) reduces to (1.8).

5. By (3.2) we have

$$\begin{aligned} \sum_{n=0}^{\infty} \sigma_n x^n &= \frac{3+gx^2}{1+gx^2-x^3} \\ &= (3+gx^2)\sum_{k=0}^{\infty} (-1)^k x^{2k}(g-x)^k \\ &= (3+gx^2)\sum_{k=0}^{\infty} (-1)^k \sum_{j=0}^{k} (-1)^j \binom{k}{j} g^{k-j}x^{j+2k} \\ &= (3+gx^2)\sum_{n=0}^{\infty} x^n \sum_{n/3\le k\le n/2} (-1)^{n+k}\binom{k}{n-2k} g^{3k-n}. \end{aligned}$$

It follows that

$$\sigma_n = \sum_{n/3 \leqq k \leqq n/2} (-1)^{n+k} \frac{n}{k} \binom{k}{n-2k} g^{3k-n} \qquad (n > 0). \tag{5.1}$$

Similarly, by (3.3),

$$\begin{aligned}\sum_{n=0}^{\infty} \sigma_{-n} x^n &= \frac{3-2gx}{1-gx-x^3} \\ &= (3-2gx) \sum_{k=0}^{\infty} x^k \sum_{j=0}^{k} \binom{k}{j} g^{k-j} x^{2j} \\ &= (3-2gx) \sum_{n=0}^{\infty} x^n \sum_{3j \leqq n} \binom{n-2j}{j} g^{n-3j}\end{aligned}$$

and we find that

$$\sigma_{-n} = \sum_{3j \leqq n} \frac{n}{n-2j} \binom{n-2j}{j} g^{n-3j} \qquad (n > 0). \tag{5.2}$$

We may now substitute from (5.1) and (5.2) in the formulas of § 3 to obtain a variety of combinatorial identities. In particular, substituting in (3.6) we get

$$\begin{aligned}&\left\{ \sum_{n/3 \leqq k \leqq n/2} (-1)^{n+k} \frac{n}{k} \binom{k}{n-2k} g^{3k-n} \right\}^2 \\ &\quad = \sum_{2n/3 \leqq k \leqq n} (-1)^k \frac{2n}{k} \binom{k}{2n-2k} g^{3k-2n} \\ &\quad\quad + 2 \sum_{3j \leqq n} \frac{n}{n-2j} \binom{n-2j}{j} g^{n-3j} \qquad (n > 0)\end{aligned} \tag{5.3}$$

and

$$\begin{aligned}&\left\{ \sum_{3j \leqq n} \frac{n}{n-2j} \binom{n-2j}{j} g^{n-3j} \right\}^2 \\ &\quad = \sum_{3j \leqq 2n} \frac{n}{n-j} \binom{2n-2j}{j} g^{2n-3j} \\ &\quad\quad + 2 \sum_{n/3 \leqq k \leqq n/2} (-1)^{n+k} \frac{n}{k} \binom{k}{n-2k} g^{3k-n} \qquad (n > 0).\end{aligned} \tag{5.4}$$

In substituting in (3.8) or (3.9) there are a number of possibilities depending on the relative sign of m and n. However we shall not take the space to write out the resulting identities.

6. Equating coefficients of powers of g in any of the polynomial identities leads to certain binomial identities. For example (5.3) gives

$$\sum_{r+s=k} \frac{n}{r}\binom{r}{n-2r}\frac{n}{s}\binom{s}{n-2s} = \frac{2n}{k}\binom{k}{2n-2k} + (-1)^k \frac{2n}{2k-n}\binom{2k-n}{n-k} \qquad \left(\frac{2n}{3} \leqq k \leqq n\right), \tag{6.1}$$

while (5.4) yields

$$\sum_{r+s=k} \frac{n}{n-2r}\binom{n-2r}{r}\frac{n}{n-2s}\binom{n-2s}{s} = \frac{n}{n-k}\binom{2n-2k}{k} + (-1)^{n+k}\frac{n}{n-k}\binom{n-k}{2n-3k} \qquad (3k \leqq 2n). \tag{6.2}$$

Binomial coefficient summations such as (6.1) and (6.2) may, if we prefer, be written in the notation of generalized hypergeometric functions. For example the left hand side of (6.1) is equal to the well-poised sum

$${}_6F_5\left[\begin{matrix} -k, \frac{n}{2}-k, \frac{n}{2}-k+\frac{1}{2}, -\frac{n}{3}, \frac{-n+1}{3}, \frac{-n+2}{3}; \\ \frac{-n+2}{2}, \frac{-n+1}{2}, \frac{n+3}{3}-k, \frac{n+2}{3}-k, \frac{n+1}{3}-k; \quad 1 \end{matrix}\right]$$

REFERENCES

[1] R. ASKEY, *Orthogonal Polynomials and Special Functions*, Society for Industrial and Applied Mathematics, Philadelphia, 1975.

[2] L. BERNSTEIN, *Zeros of combinatorial functions and combinatorial identities*, Houston J. Math., 2 (1976), pp. 9–16.

[3] ———, *Zeros of the functions* $f(n)=\sum_{i=0}(-1)^i\binom{n-2i}{i}$, J. Number Theory, 6 (1974), pp. 264–270.

[4] G. SZEGÖ, *Orthogonal Polynomials*, American Mathematical Society, Providence, RI, 1939.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

PRODUCT FORMULAS AND NICHOLSON-TYPE INTEGRALS FOR JACOBI FUNCTIONS. I: SUMMARY OF RESULTS*

LOYAL DURAND†

Abstract. Nicholson's formula gives a generalization of the relation $\sin^2 x + \cos^2 x = 1$ to the case of Bessel functions. We present a similar result which relates the sum of squares of the Jacobi functions $P_n^{(\alpha,\beta)}(x)$ and $Q_n^{(\alpha,\beta)}(x)$ to an integral over a single Jacobi function of the second kind, with the integrand positive. The Nicholson-type formula is a special case of a general product formula for two Jacobi functions of the second kind with different arguments, $Q_n^{(\alpha,\beta)}(z_1)Q_n^{(\alpha,\beta)}(z_2)$. Various confluent limits of these expressions give Nicholson-type integrals and product formulas for general Gegenbauer, Laguerre, Bessel, and Hermite functions. These results are summarized in the present paper. Derivations and applications will be given elsewhere.

1. Introduction. In 1910, J. W. Nicholson [1] gave a generalization the familiar relation $e^{ix}e^{-ix} = \sin^2 x + \cos^2 x = 1$ for the case of Bessel functions. Nicholson's result expresses the sum $H_\nu^{(1)}H_\nu^{(2)} = J_\nu^2 + Y_\nu^2$ as an integral over a hyperbolic Bessel function with the integrand positive,

$$J_\nu^2(x) + Y_\nu^2(x) = \frac{8}{\pi^2}\int_0^\infty K_0(2x \sinh t)\cosh 2\nu t\, dt. \tag{1.1}$$

Derivations of (1.1) and some related integrals are given in Watson [2]. Nicholson's result has been of considerable importance in the theory of Bessel functions. It was used by Nicholson [1] to obtain asymptotic expansions of $J_\nu^2 + Y_\nu^2$ for large x, and by Watson [2] to derive bounds on the Bessel functions and to establish a number of interesting monotonicity properties. It follows, for example, from the analysis given by Watson [2, § 13.74] that the function $x[J_\nu^2(x) + Y_\nu^2(x)]$ is a completely monotonic function of x for $\nu > \frac{1}{2}$, a result used by Lorch and Szego [3] to prove a number of remarkable monotonicity properties of the nth differences of the zeros of Bessel functions and the areas under successive arches.

One would expect expressions analogous to Nicholson's integral to exist for the classical orthogonal polynomials, as these reduce to Bessel functions in appropriate confluent limits. However, despite the extensive literature on these polynomials [4], [5], [6], no such results were known until 1971 when the author [7] derived a Nicholson-type formula for Gegenbauer functions of arbitrary degree and order. I have recently obtained the corresponding results for general Jacobi functions.

Our basic result is a formula which expresses a product of two Jacobi functions of the second kind, $Q_n^{(\alpha,\beta)}(z_1)Q_n^{(\alpha,\beta)}(z_2)$, as an integral involving a third function $Q_n^{(\alpha,\beta)}(Z)$ with a modified argument. The form of this expression was suggested by the integrated form of Koornwinder's addition theorem for the

* Received by the editors February 12, 1976.

† Department of Physics, University of Wisconsin, Madison, Wisconsin 53706. This work was supported in part by the University of Wisconsin Research Committee with funds granted by the Wisconsin Alumni Research Foundation, and in part by the U.S. Energy Research and Development Administration.

Jacobi polynomials [8]. A special case of the product formula gives the Nicholson-type integral for Jacobi functions. By considering various confluent limits of the Jacobi functions, we obtain analogues of the product formula and Nicholson's integral for Gegenbauer, Laguerre, Bessel, and Hermite functions. These results will be summarized here. Details of the derivations and some applications involving bounds, addition theorems, asymptotic expansions, and higher monotonicity properties of the functions will be presented in a series of papers in preparation.

2. Results for Jacobi functions.

2.1. Definitions. Let $P_n^{(\alpha,\beta)}(z)$ and $Q_n^{(\alpha,\beta)}(z)$ be the Jacobi functions of the first and second kind defined for arbitrary α, β, n and complex z by [4, § 4.21, 4.61]

$$P_n^{(\alpha,\beta)}(z)=\frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)\Gamma(\alpha+1)}{}_2F_1\left(-n,n+\alpha+\beta+1;\alpha+1;\frac{1-z}{2}\right), \tag{2.1}$$

$$\begin{aligned}Q_n^{(\alpha,\beta)}(z)&=2^{n+\alpha+\beta}\frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{\Gamma(2n+\alpha+\beta+2)}(z-1)^{-n-\alpha-1}(z+1)^{-\beta}\\&\quad\cdot {}_2F_1\left(n+1,n+\alpha+1;2n+\alpha+\beta+2;\frac{2}{1-z}\right).\end{aligned} \tag{2.2}$$

We define the Jacobi functions for real argument x "on the cut", $-1<x<1$, by[1]

$$\begin{aligned}P_n^{(\alpha,\beta)}(x)&=\frac{i}{\pi}[e^{i\pi\alpha}Q_n^{(\alpha,\beta)}(x+i0)-e^{-i\pi\alpha}Q_n^{(\alpha,\beta)}(x-i0)]\\&=P_n^{(\alpha,\beta)}(x\pm i0), \qquad -1<x\leqq 1,\end{aligned} \tag{2.3}$$

$$Q_n^{(\alpha,\beta)}(x)=\tfrac{1}{2}[e^{i\pi\alpha}Q_n^{(\alpha,\beta)}(x+i0)+e^{-i\pi\alpha}Q_n^{(\alpha,\beta)}(x-i0)], \qquad -1<x<1. \tag{2.4}$$

$P_n^{(\alpha,\beta)}(x)$ and $Q_n^{(\alpha,\beta)}(x)$ are real for real n, α, β.

[1] The definition of the (seldom-used) function $Q_n^{(\alpha,\beta)}(x)$ given in (2.4) differs from that given in [6, § 10.8(22)] and [4, § 4.62.9],

$$Q_n^{(\alpha,\beta)}(x)=\tfrac{1}{2}[Q_n^{(\alpha,\beta)}(x+i0)+Q_n^{(\alpha,\beta)}(x-i0)].$$

The latter definition destroys the analogy between $P_n^{(\alpha,\beta)}(\cos\theta)$, $Q_n^{(\alpha,\beta)}(\cos\theta)$, and the trigonometric functions, and is not appropriate for our purposes. With our definition (2.4), the functions $Q_n^{(\alpha,\beta)}(\cos\theta\pm i0)$ are the analogues for Jacobi functions of the complex exponentials $e^{\pm i\theta}$. In fact, for large n,

$$Q_n^{(\alpha,\beta)}(\cos\theta\pm i0)\sim\frac{1}{2}\left(\frac{\pi}{n}\right)^{1/2}\left(\sin\frac{\theta}{2}\right)^{-\alpha-1/2}\left(\cos\frac{\theta}{2}\right)^{-\beta-1/2}e^{\mp iN\theta\mp i(\alpha+1/2)\pi/2}, \qquad N=n+\frac{\alpha}{2}+\frac{\beta}{2}+\frac{1}{2}.$$

2.2. The product formula. Our basic result is an expression for the product of two Jacobi functions of the second kind,

$$
\begin{aligned}
&[(z_1-1)(z_2-1)]^{l+m/2}[(z_1+1)(z_2+1)]^{m/2}Q_{n-l-m}^{(\alpha+2l+m,\beta+m)}(z_1)Q_{n-l-m}^{(\alpha+2l+m,\beta+m)}(z_2)\\
&\qquad = N_{n,l,m}^{\alpha,\beta}\int_1^\infty dr\int_1^\infty dt\, Q_n^{(\alpha,\beta)}(Z)P_l^{(\alpha-\beta-1,\beta+m)}(2r^2-1)\\
&\qquad\quad \cdot C_m^\beta(t)(t^2-1)^{\beta-1/2}(r^2-1)^{\alpha-\beta-1}r^{2\beta+m+1}.
\end{aligned}
\tag{2.5}
$$

Here $C_m^\beta(t)$ is a Gegenbauer function of the first kind, defined in (3.1). The argument of the Jacobi function $Q_n^{(\alpha,\beta)}(Z)$ on the right hand side of (2.5) is given by

$$
Z=z_1z_2+[(z_1^2-1)(z_2^2-1)]^{1/2}rt+\tfrac{1}{2}(z_1-1)(z_2-1)(r^2-1),
\tag{2.6}
$$

and the normalization factor $N_{n,l,m}^{\alpha,\beta}$ is given by

$$
\begin{aligned}
N_{n,l,m}^{\alpha,\beta} &= 2^{2\beta+2l+2m}\frac{\Gamma(\beta)\Gamma(n+\beta-l+1)\Gamma(n+\alpha+l+1)}{\Gamma(n-l-m+1)\Gamma(n+\beta+1)\Gamma(n+\alpha+\beta+l+m+1)}\\
&\quad\cdot\frac{\Gamma(n+\alpha+\beta+1)\Gamma(m+1)\Gamma(l+1)}{\Gamma(m+2\beta)\Gamma(l+\alpha-\beta)}.
\end{aligned}
\tag{2.7}
$$

The expression (2.5) holds for complex n, α, β, l, m with $\operatorname{Re}\alpha>\operatorname{Re}\beta>-\frac{1}{2}$, $\operatorname{Re}(m+\beta)\geqq 0$, $\operatorname{Re}(l+\alpha/2+m/2)\geqq 0$, $\operatorname{Re}(n-l-m+1)>0$, and $\operatorname{Re}(n+\alpha-\beta-m+1)>0$. The Q's are holomorphic in the complex plane cut from $+1$ to $-\infty$. Equation (2.5) holds in the form given for $|\arg(z_1\pm 1)|<\pi$, $|\arg(z_2\pm 1)|<\pi$, $|\arg(z_1-1)(z_2-1)|<\pi$, $|\arg\sqrt{z_1^2-1}\sqrt{z_2^2-1}|<\pi$, and can be continued elsewhere by using the reflection symmetries of the Q's for $z\to e^{\pm i\pi}z$. For $l=m=0$, (2.5) reduces to a relatively simple expression,

$$
\begin{aligned}
Q_n^{(\alpha,\beta)}(z_1)Q_n^{(\alpha,\beta)}(z_2) &= 2\sqrt{\pi}\frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)\Gamma(\beta+\frac{1}{2})\Gamma(\alpha-\beta)}\\
&\quad\cdot\int_1^\infty dr\int_1^\infty dt\, Q_n^{(\alpha,\beta)}(Z)\,(t^2-1)^{\beta-1/2}(r^2-1)^{\alpha-\beta-1}r^{2\beta+1},
\end{aligned}
\tag{2.8}
$$

$$
\operatorname{Re}\alpha>\operatorname{Re}\beta>\tfrac{1}{2},\quad \operatorname{Re}(n+\beta+1)>0,\quad \operatorname{Re}(n+\alpha-\beta+1)>0.
$$

The double integral in (2.8) can be converted to a single integral by making the change of variables

$$
\begin{aligned}
e^{\psi}\cosh t_3 &= \cosh t_1\cosh t_2+\sinh t_1\sinh t_2 r e^{\phi},\\
e^{-\psi}\cosh t_3 &= \cosh t_1\cosh t_2+\sinh t_1\sinh t_2 r e^{-\phi},
\end{aligned}
\tag{2.9}
$$

with $z_1=\cosh 2t_1$, $z_2=\cosh 2t_2$, $t=\cosh\phi$, and integrating over ψ. This gives an expression analogous to the kernel form of the expression for the product of two

Jacobi polynomials derived by Gasper [9] and Koornwinder [10],

$$Q_n^{(\alpha,\beta)}(\cosh 2t_1)Q_n^{(\alpha,\beta)}(\cosh 2t_2)=\sqrt{\pi}\,2^{\alpha-1/2}\frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)\Gamma(\alpha+\frac{1}{2})}$$

(2.10) $$\cdot(\sinh t_1\sinh t_2)^{-2\alpha}(\cosh t_1\cosh t_2)^{\alpha-\beta-1}\int_0^\infty Q_n^{(\alpha,\beta)}(\cosh 2t_3)$$

$$\cdot(B^2-1)^{\alpha-1/2}{}_2F_1\left(\alpha+\beta,\alpha-\beta;\alpha+\frac{1}{2};\frac{1-B}{2}\right)(\cosh t_3)^{\alpha+\beta}\sinh t_3\,dt_3,$$

where

(2.11) $$B=\frac{\cosh^2 t_1+\cosh^2 t_2+\cosh^2 t_3-1}{2\cosh t_1\cosh t_2\cosh t_3}.$$

2.3. Nicholson-type integrals. We obtain Nicholson-type integrals for Jacobi functions by letting z_1 and z_2 in (2.5) approach a real point x on opposite sides of the cut, with $-1<x<1$. If we use the relation

(2.12) $$Q_n^{(\alpha,\beta)}(x+i0)Q_n^{(\alpha,\beta)}(x-i0)=[Q_n^{(\alpha,\beta)}(x)]^2+\frac{\pi^2}{4}[P_n^{(\alpha,\beta)}(x)]^2$$

which follows from (2.3) and (2.4), we find from (2.8) that

$$[Q_n^{(\alpha,\beta)}]^2+\frac{\pi^2}{4}[P_n^{(\alpha,\beta)}(x)]^2=2\sqrt{\pi}\frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)\Gamma(\beta+\frac{1}{2})\Gamma(\alpha-\beta)}$$

(2.13) $$\cdot\int_1^\infty dr\int_1^\infty dt\,Q_n^{(\alpha,\beta)}(x^2+(1-x^2)rt+\tfrac{1}{2}(1-x)^2(r^2-1))$$

$$\cdot(t^2-1)^{\beta-1/2}(r^2-1)^{\alpha-\beta-1}r^{2\beta-1},\qquad -1<x<1.$$

The integrand is real and positive for α, β, n real. A more general result follows in a similar fashion from (2.5),

$$(1-x)^{2l+m}(1+x)^m\left\{[Q_{n-l-m}^{(\alpha+2l+m,\beta+m)}(x)]^2+\frac{\pi^2}{4}[P_{n-l-m}^{(\alpha+2l+m,\beta+m)}(x)]^2\right\}$$

(2.14) $$=N_{n,l,m}^{\alpha,\beta}\int_1^\infty dr\int_1^\infty dt\,Q_n^{(\alpha,\beta)}(x^2+(1-x^2)rt+\tfrac{1}{2}(1-x)^2(r^2-1))$$

$$\cdot P_l^{(\alpha-\beta-1,\beta+m)}(2r^2-1)C_m^\beta(t)(t^2-1)^{\beta-1/2}(r^2-1)^{\alpha-\beta-1}r^{2\beta+m+1}.$$

2.4. Laplace-type integral representation. We can also obtain an interesting new integral representation for $Q_n^{(\alpha,\beta)}(z)$ from (2.8) by letting $z_1=z$ and taking $z_2\to\infty$. Comparison of the two sides of the equation gives

$$Q_n^{(\alpha,\beta)}(z)=2\sqrt{\pi}\frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)\Gamma(\beta+\frac{1}{2})\Gamma(\alpha-\beta)}$$

(2.15) $$\cdot\int_1^\infty dr\int_1^\infty dt\,[z+(z^2-1)^{1/2}rt+\tfrac{1}{2}(z-1)(r^2-1)]^{-n-\alpha-\beta-1}$$

$$\cdot(t^2-1)^{\beta-1/2}(r^2-1)^{\alpha-\beta-1}r^{2\beta+1}.$$

This integral representation is similar in structure to the Laplace-type integral representation for the Jacobi polynomials $P_n^{(\alpha,\beta)}(z)$, $n=$ integer, derived by Koornwinder [8].

3. Results for Gegenbauer functions.

3.1. Definitions. The Gegenbauer functions of the first kind, $C_n^\alpha(z)$, are defined in [6, § 3.15],

$$\begin{aligned} C_n^\alpha(z) &= \frac{\Gamma(n+2\alpha)}{\Gamma(2\alpha)\Gamma(n+1)}\,{}_2F_1\Big(-n, n+2\alpha; \alpha+\frac{1}{2}; \frac{1-z}{2}\Big) \\ &= 2^{-2\alpha+1}\sqrt{\pi}\frac{\Gamma(n+2\alpha)}{\Gamma(\alpha)\Gamma(n+\alpha+\frac{1}{2})}P_n^{(\alpha-1/2,\alpha-1/2)}(z). \end{aligned} \tag{3.1}$$

The functions of the second kind, $D_n^\alpha(z)$, will be defined as in [7],[2]

$$\begin{aligned} D_n^\alpha(z) &= e^{i\pi\alpha}\frac{\Gamma(n+2\alpha)}{\Gamma(\alpha)\Gamma(n+\alpha+1)}(2z)^{-n-2\alpha}\,{}_2F_1\Big(\frac{n}{2}+\alpha, \frac{n}{2}+\alpha+\frac{1}{2}; n+\alpha+1; z^{-2}\Big) \\ &= 2^{-2\alpha+1}\,e^{i\pi\alpha}\frac{1}{\sqrt{\pi}}\frac{\Gamma(n+2\alpha)}{\Gamma(\alpha)\Gamma(n+\alpha+1)}Q_n^{(\alpha-1/2,\alpha-1/2)}(z). \end{aligned} \tag{3.2}$$

The phase factor $e^{i\pi\alpha}$ is included in the definition of $D_n^\alpha(z)$ so that D_n^α and C_n^α satisfy the same recurrence relations with respect to α [7] ($P_n^{(\alpha,\beta)}$ and $Q_n^{(\alpha,\beta)}$ as usually defined satisfy different recurrence relations with respect to α).

We define the Gegenbauer functions for real argument x "on the cut", $-1<x<1$, by

$$C_n^\alpha(x) = D_n^\alpha(x+i0)+e^{-2\pi i\alpha}D_n^\alpha(x-i0) = C_n^\alpha(x\pm i0), \qquad -1<x\leqq 1, \tag{3.3}$$

$$D_n^\alpha(x) = -iD_n^\alpha(x+i0)+i\,e^{-2\pi i\alpha}D_n^\alpha(x-i0), \qquad -1<x<1. \tag{3.4}$$

$C_n^\alpha(x)$ and $D_n^\alpha(x)$ are real for real n and α.

3.2. The product formula. The product formula for the Gegenbauer functions of the second kind can be derived from (2.5) by converting the integral on r into a contour integral on a contour around the segment of the real axis $1\leqq r<\infty$, and then letting β equal α. The integrand has a first order pole at $r=1$ for $\alpha=\beta$, and the integration over r is trivial. The result is nonzero only for $l=0$. After using the definition (3.2), we obtain the product formula [7][3]

$$\begin{aligned} &(z_1^2-1)^{m/2}(z_2^2-1)^{m/2}D_{n-m}^{\alpha+m}(z_1)D_{n-m}^{\alpha+m}(z_2) \\ &\quad = 2^{-2\alpha-2m+1}\frac{\Gamma(2\alpha-1)\Gamma(m+1)\Gamma(n+2\alpha+m)}{[\Gamma(\alpha+m)]^2\Gamma(2\alpha+m-1)\Gamma(n-m+1)} \\ &\quad\quad \cdot e^{i\pi(\alpha+2m)}\int_1^\infty D_n^\alpha(Z)C_m^{\alpha-1/2}(t)\,(t^2-1)^{\alpha-1}\,dt, \end{aligned} \tag{3.5}$$

[2] The functions $D_n^\alpha(z)$ defined above differ from those of Robin [5, § 170(93)] by a factor $\pi^{-1}e^{2\pi i(\alpha-1/4)}$ which simplifies the connections between the C's and D's. The definition of $D_n^\alpha(z)$ given in [6, p. 175, corrections on p. 2] is inappropriate for our purposes.

[3] The phase factor on the right hand side of (3.5) is given incorrectly in [7, Eq. (13)] as $e^{i\pi\alpha}$.

where

$$Z = z_1 z_2 + (z_1^2 - 1)^{1/2}(z_2^2 - 1)^{1/2} t. \tag{3.6}$$

This result holds for complex n, α, and m with $\operatorname{Re}\alpha > 0$, $\operatorname{Re}(\alpha + m) \geqq 0$, and $\operatorname{Re}(n - m + 1) > 0$. The D's are holomorphic in the complex plane cut from $+1$ to $-\infty$. The expression in (3.5) is valid in the form given for $|\arg(z_1 \pm 1)| < \pi$, $|\arg(z_2 \pm 1)| < \pi$, $|\arg\sqrt{z_1^2 - 1}\sqrt{z_2^2 - 1}| < \pi$, and can be continued to other regions by the use of the reflection symmetry of the D's for $z \to e^{\pm i\pi} z$,

$$D_n^\alpha(z\, e^{\pm i\pi}) = e^{\mp i\pi(n + 2\alpha)} D_n^\alpha(z). \tag{3.7}$$

An alternative derivation of (3.5) based on the addition formula for Gegenbauer functions is given in [7].

For $\alpha = \frac{1}{2}$ and $m = 0$, (3.5) gives a product formula for the Legendre functions $Q_n(z)$,

$$D_n^{1/2}(z) = \frac{i}{\pi} Q_n(z), \tag{3.8}$$

$$Q_n(z_1) Q_n(z_2) = \int_1^\infty Q_n(z_1 z_2 + (z_1^2 - 1)^{1/2}(z_2^2 - 1)^{1/2} t)(t^2 - 1)^{-1/2}\, dt, \tag{3.9}$$

$$\operatorname{Re} n > -1, \qquad |\arg[(z_1^2 - 1)^{1/2}(z_2^2 - 1)^{1/2}]| < \pi.$$

3.3. Nicholson-type integrals. We obtain Nicholson-type integrals for the Gegenbauer functions by letting z_1 and z_2 in (3.5) approach a real point, x, $-1 < x < 1$, on opposite sides of the cut. If we use the relation

$$D_n^\alpha(x + i0) D_n^\alpha(x - i0) = \tfrac{1}{4} e^{2\pi i\alpha}\{[C_n^\alpha(x)]^2 + [D_n^\alpha(x)]^2\} \tag{3.10}$$

which follows from (3.3) and (3.4), we find from (3.5) that

$$\begin{aligned} &(1 - x^2)^m \{[C_{n-m}^{\alpha+m}(x)]^2 + [D_{n-m}^{\alpha+m}(x)]^2\} \\ &\quad = 2^{-2\alpha - 2m + 3} \frac{\Gamma(2\alpha - 1)\Gamma(m + 1)\Gamma(n + 2\alpha + m)}{[\Gamma(\alpha + m)]^2 \Gamma(2\alpha + m - 1)\Gamma(n - m + 1)} \\ &\quad \cdot e^{-i\pi\alpha} \int_1^\infty D_n^\alpha(x^2 + (1 - x^2)t) C_m^{\alpha - 1/2}(t)(t^2 - 1)^{\alpha - 1}\, dt, \qquad -1 < x < 1, \end{aligned} \tag{3.11}$$

where the functions on the left hand side of the expression are the Gegenbauer functions "on-the cut". For $m = 0$ and $\alpha = \frac{1}{2}$, we obtain a Nicholson-type formula for the ordinary Legendre functions,

$$[P_n(x)]^2 + \frac{4}{\pi^2}[Q_n(x)]^2 = \frac{4}{\pi^2} \int_1^\infty Q_n(x^2 + (1 - x^2)t)(t^2 - 1)^{-1/2}\, dt, \qquad -1 < x < 1. \tag{3.12}$$

3.4. Laplace-type integral representation. We obtain an integral representation for $D_n^\alpha(z)$ similar in structure to the Laplace integral for $C_n^\alpha(z)$ by letting $z_1 = z$ in (3.5) and taking $z_2 \to \infty$. Comparison of the asymptotic forms of the two sides of the equation gives

$$(z^2-1)^{m/2}D_{n-m}^{\alpha+m}(z)=2^{-m-2\alpha+1}\frac{\Gamma(2\alpha-1)\Gamma(n+2\alpha)\Gamma(m+1)}{\Gamma(\alpha)\Gamma(2\alpha+m-1)\Gamma(n-m+1)\Gamma(\alpha+m)}$$

$$\cdot e^{i\pi(\alpha+m)}\int_1^\infty [z+(z^2-1)^{1/2}t]^{-n-2\alpha}C_m^{\alpha-1/2}(t) \cdot (t^2-1)^{\alpha-1}\,dt, \tag{3.13}$$

$$\operatorname{Re}\alpha>0, \qquad \operatorname{Re}(n-m+1)>0.$$

This reduces to a known result for the associated Legendre functions for $m=0$ [6, § 3.7(2)].

4. Results for Laguerre functions.

4.1. Definitions. The Laguerre functions $L_n^\alpha(z)$ for general values of n and α are defined in terms of confluent hypergeometric functions [6, § 6.9.2(37)],

$$L_n^\alpha(z)=\frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)\Gamma(\alpha+1)}\Phi(-n,\alpha+1,z) = \frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)\Gamma(\alpha+1)}\,{}_1F_1(-n;\alpha+1;z). \tag{4.1}$$

A second solution to the Laguerre equation is given by [6, § 6.7]

$$N_n^\alpha(z)=\tfrac{1}{2}\Gamma(n+\alpha+1)\,e^z\Psi(n+\alpha+1,\alpha+1,-z). \tag{4.2}$$

We will define the principal branch of the many valued function $N_n^\alpha(z)$ by the condition $0<\arg z<2\pi$, with $-z=e^{-i\pi}z$. With these choices for the functions of the first and second kind, $L_n^\alpha(z)$ and $N_n^\alpha(z)$ are confluent limits of the Jacobi functions,

$$L_n^\alpha(z)=\lim_{\beta\to\infty}P_n^{(\alpha,\beta)}\left(1+e^{-i\pi}\frac{2z}{\beta}\right), \tag{4.3}$$

$$N_n^\alpha(z)=\lim_{\beta\to\infty}Q_n^{(\alpha,\beta)}\left(1+e^{-i\pi}\frac{2z}{\beta}\right), \qquad 0<\arg z<2\pi. \tag{4.4}$$

The Laguerre functions on the cut $0\leqq z<\infty$ will be defined as

$$L_n^\alpha(x)=\frac{i}{\pi}[e^{i\pi\alpha}N_n^\alpha(x-i0)-e^{-i\pi\alpha}N_n^\alpha(x+i0)], \tag{4.5}$$

$$N_n^\alpha(x)=\tfrac{1}{2}[e^{i\pi\alpha}N_n^\alpha(x-i0)+e^{-i\pi\alpha}N_n^\alpha(x+i0)], \tag{4.6}$$

and correspond to confluent limits of the Jacobi functions on the cut,

$$L_n^\alpha(x)=\lim_{\beta\to\infty}P_n^{(\alpha,\beta)}\left(1-\frac{2x}{\beta}\right), \tag{4.7}$$

$$N_n^\alpha(x)=\lim_{\beta\to\infty}Q_n^{(\alpha,\beta)}\left(1-\frac{2x}{\beta}\right), \qquad 0\leqq x<\infty. \tag{4.8}$$

4.2. A product formula. We obtain a product formula for the Laguerre functions of the second kind from (2.10) by making a change of variables

$$\cosh 2t_3 = \cosh 2t_1 \cosh 2t_2 + \sinh 2t_1 \sinh 2t_2 \cosh t, \tag{4.9}$$

replacing $\cosh 2t_1$ and $\cosh 2t_2$ by $1+e^{-i\pi}(2z_1/\beta)$ and $1+e^{-i\pi}(2z_2/\beta)$, and considering the limit $\beta \to \infty$. Use of the confluence relation (4.4) gives the product formula

$$\begin{aligned} N_n^\alpha(z_1)N_n^\alpha(z_2) &= \sqrt{\pi}\, 2^{\alpha-1/2}\frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)} \\ &\cdot \int_0^\infty N_n^\alpha\left((z_2^{1/2}+z_2^{1/2})^2+4(z_1z_2)^{1/2}\sinh^2\frac{t}{2}\right) e^{-(z_1z_2)^{1/2}\cosh t} \\ &\cdot I_{\alpha-1/2}((z_1z_2)^{1/2}\sinh t)[(z_1z_2)^{1/2}\sinh t]^{-\alpha+1/2}(\sinh t)^{2\alpha}\,dt, \end{aligned} \tag{4.10}$$

$$0<\arg z_1<2\pi, \quad 0<\arg z_2<2\pi, \quad \tfrac{1}{2}\pi<\arg(z_1z_2)^{1/2}<\tfrac{3}{2}\pi,$$
$$\operatorname{Re}\alpha>0, \quad \operatorname{Re}(n+1)>0,$$

where $I_{\alpha-1/2}(z)$ is a hyperbolic Bessel function [2, § 3.7]. This result is the analogue for the Laguerre functions of the second kind of the product formula for $L_n^\alpha(z_1)L_n^\alpha(z_2)$ derived by Watson [11].

4.3. Nicholson-type integral. We obtain a Nicholson-type integral for the Laguerre functions by letting the variables z_1 and z_2 in (4.10) approach a real point x, $0<x<\infty$, from opposite sides of the positive real axis. If we use the relation

$$N_n^\alpha(x+i0)N_n^\alpha(x-i0) = [N_n^\alpha(x)]^2+\frac{\pi^2}{4}[L_n^\alpha(x)]^2 \tag{4.11}$$

which follows from (4.5) and (4.6), we find that

$$\begin{aligned} [N_n^\alpha(x)]^2+\frac{\pi^2}{4}[L_n^\alpha(x)]^2 &= \sqrt{\pi}\left(\frac{2}{x}\right)^{\alpha-1/2}\frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)} \\ &\cdot \int_0^\infty N_n^\alpha\left(-4x\sinh^2\frac{t}{2}\right) e^{x\cosh t} \\ &\cdot I_{\alpha-1/2}(x\sinh t)(\sinh t)^{\alpha+1/2}\,dt, \qquad 0<x<\infty, \end{aligned} \tag{4.12}$$

where the functions on the left hand side of (4.12) are the Laguerre functions "on the cut". The integrand in (4.12) is real and positive for n and α real.

5. Results for Bessel functions.

5.1. Product formulas. We can derive product formulas for the hyperbolic Bessel functions $K_\nu(z)$ [2, § 3.7] from our product formula for the Gegenbauer functions, (3.5), by using a confluent limit of the D's [7],

$$\lim_{n\to\infty} n^{-2\alpha+1}e^{-i\pi\alpha}D_n^\alpha\left(1+\frac{z^2}{2n^2}\right) = \frac{1}{\sqrt{\pi}\Gamma(\alpha)}(2z)^{-\alpha+1/2}K_{\alpha-1/2}(z). \tag{5.1}$$

Thus, if we let $\nu=\alpha-\frac{1}{2}$ and $t=\cosh\phi$ in (3.5), replace z_1 and z_2 by $1+\frac{1}{2}(z_1/n)^2$ and $1+\frac{1}{2}(z_2/n)^2$, and let $n\to\infty$, we find that [7][4]

$$\frac{K_{\nu+m}(z_1)}{z_1^\nu}\frac{K_{\nu+m}(z_2)}{z_2^\nu}=2^{\nu-1}\frac{\Gamma(\nu)\Gamma(m+1)}{\Gamma(m+2\nu)} \cdot\int_0^\infty \omega^{-\nu}K_\nu(\omega)C_m^\nu(\cosh\phi)(\sinh\phi)^{2\nu}\,d\phi, \tag{5.2}$$

where

$$\omega=(z_1^2+z_2^2+2z_1z_2\cosh\phi)^{1/2},\qquad |\arg\omega|<\pi/2. \tag{5.3}$$

This expression is similar in structure to the product formulas for ordinary Bessel functions given by Gegenbauer [2, § 11.42].

5.2. Nicholson-type integrals. We obtain a generalization of Nicholson's integral from (5.2) by replacing z_1 and z_2 in (5.2) by $xe^{i(\pi/2)}$ and $xe^{-i(\pi/2)}$, $0<x<\infty$, and using the relations

$$K_\nu(xe^{\pm i(\pi/2)})=\mp i\frac{\pi}{2}e^{\pm i\nu\pi/2}[J_\nu(x)\mp iY_\nu(x)]. \tag{5.4}$$

After a change of variables $\phi\to 2t$, we obtain

$$J_{\nu+m}^2(x)+Y_{\nu+m}^2(x)=\frac{4}{\pi^2}\frac{\Gamma(\nu)\Gamma(m+1)}{\Gamma(m+2\nu)}(4x)^\nu \cdot\int_0^\infty K_\nu(2x\sinh t)C_m^\nu(\cosh 2t) \cdot(\cosh t)^{2\nu}(\sinh t)^\nu\,dt. \tag{5.5}$$

We recover Nicholson's formula (1.1) by taking the limit $\nu\to 0$, m arbitrary, and using the relation

$$\lim_{\nu\to 0}\frac{\Gamma(\nu)\Gamma(m+1)}{\Gamma(m+2\nu)}C_m^\nu(\cosh 2t)=2\cosh 2mt. \tag{5.6}$$

6. Results for Hermite functions.

6.1. Definitions. We define the Hermite functions $H_n(x)$ and $G_n(x)$ for arbitrary complex n in terms of confluent hypergeometric functions [6, Chap. 6],

$$\begin{aligned} H_n(x)&=2^n\Psi\left(-\frac{n}{2},\frac{1}{2},x^2\right)\\ &=\frac{2^n}{\sqrt{\pi}}\left[\cos\frac{n\pi}{2}\Gamma\left(\frac{n}{2}+\frac{1}{2}\right)\Phi\left(-\frac{n}{2},\frac{1}{2},x^2\right)\right.\\ &\qquad\left.+2x\sin\frac{n\pi}{2}\Gamma\left(\frac{n}{2}+1\right)\Phi\left(-\frac{n}{2}+\frac{1}{2},\frac{3}{2},x^2\right)\right], \end{aligned} \tag{6.1}$$

[4] The factor $2^{\nu-1}$ on the right hand side of (5.2) is given incorrectly in [7, Eq. (44)] as 2^ν.

(6.2)
$$G_n(x)=\frac{2^n}{\sqrt{\pi}}\left[-\sin\frac{n\pi}{2}\Gamma\left(\frac{n}{2}+\frac{1}{2}\right)\Phi\left(-\frac{n}{2},\frac{1}{2},x^2\right)\right.$$
$$\left.+2x\cos\frac{n\pi}{2}\Gamma\left(\frac{n}{2}+1\right)\Phi\left(-\frac{n}{2}+\frac{1}{2},\frac{3}{2},x^2\right)\right].$$

These functions are related to the standard parabolic cylinder functions $D_n(\pm z)$ [6, § 8.2] by

(6.3)
$$H_n(x)=2^{n/2}e^{x^2/2}D_n(\sqrt{2}\,x),$$

(6.4)
$$G_n(x)=2^{n/2}e^{x^2/2}\frac{1}{\sin\pi n}[\cos\pi n D_n^\alpha(\sqrt{2}\,x)-D_n(-\sqrt{2}\,x)].$$

As defined, $G_n(x)$ and $H_n(x)$ are simple confluent limits of the Gegenbauer functions C_n^α and D_n^α "on the cut",

(6.5)
$$H_n(x)=\lim_{\alpha\to\infty}\alpha^{-n/2}\Gamma(n+1)C_n^\alpha(x/\sqrt{\alpha}),$$

(6.6)
$$G_n(x)=\lim_{\alpha\to\infty}\alpha^{-n/2}\Gamma(n+1)D_n^\alpha(x/\sqrt{\alpha}).$$

The result for $H_n(x)$ is well-known [4, § 5.6(3)]; that for $G_n(x)$ is easily derived.

6.2. Nicholson-type integral. We obtain a Nicholson-type integral for the Hermite functions by considering a confluent limit of (3.11). Let $n\to n+m$, $x\to x/\sqrt{m}$, and $t=\cosh 2\phi$. The limit $m\to\infty$ then gives

(6.7)
$$e^{-x^2}[H_n^2(x)+G_n^2(x)]=\pi^{-1}2^{n+1}\Gamma(n+1)\int_0^\infty \exp[-(2n+1)\phi+x^2\tanh\phi]$$
$$\cdot(\sinh\phi\cosh\phi)^{-1/2}\,d\phi.$$

It follows immediately from this result that the expression on the left is an absolutely monotonic function of x.

Acknowledgments. The author would like to thank the members of the School of Natural Sciences at the Institute for Advanced Study and the Theoretical Division of Los Alamos Scientific Laboratory for the hospitality accorded him while parts of these results were derived and the present paper written.

REFERENCES

[1] J. W. NICHOLSON, *The Asymptotic Expansions of Bessel Functions*, Philos. Mag., 19 (1910), pp. 228–249, Eq. 34; *Notes on Bessel Functions*, Quart. J. of Pure and Appl. Math., 42 (1911), pp. 216–224, Eq. 34.

[2] G. N. WATSON, *Theory of Bessel Functions*, 2nd Edition, Cambridge University Press, Cambridge, 1966.

[3] L. LORCH AND P. SZEGO, *Higher monotonicity properties of certain Sturm–Liouville functions*, Acta Math., 109 (1963), pp. 55–73.

[4] G. SZEGÖ, *Orthogonal Polynomials*, Colloquium Publications, vol. 23, Amer. Math. Soc., Providence, RI, 1972.

[5] L. ROBIN, *Fonctions Sphérique de Legendre et Fonctions Sphéroïdales*, Gauthiers-Villars, Paris, 1959.

[6] A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, *Higher Transcendental Functions*, vols. I, II, McGraw-Hill, New York, 1953.

[7] L. Durand, *Nicholson-type integrals for products of Gegenbauer functions*, Notices Amer. Math. Soc., 20, A392 (1973); *Nicholson-type integrals for products of Gegenbauer functions and related topics*, Theory and Applications of Special Functions, R. Askey, editor, Academic Press, New York, 1975, pp. 353–374.

[8] T. Koornwinder, *The addition theorem for Jacobi polynomials, I. Summary of results*, Indag. Math., 34 (1972), pp. 188–191.

[9] G. Gasper, *Positivity and the convolution structure of Jacobi series*, Ann. of Math., 93 (1971), pp. 112–118.

[10] T. Koornwinder, *Jacobi polynomials, II. An analytic proof of the product formula*, this Journal, 5 (1974), pp. 125–137, § 5.

[11] G. N. Watson, *Another note on Laguerre polynomials*, J. London Math. Soc., 14 (1939), pp. 19–22.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

CALCULATION OF SOME EXTREMAL CONFORMAL MAPPINGS*

E. GRASSMANN† AND J. ROKNE†

Abstract. The following two extremum problems are treated:

(i) To find a continuum that connects n given points in the complex plane $\mathbb{C}$ and has minimal capacity;

(ii) To find a doubly connected domain D that separates two given finite sets of points in the complex plane and has maximum modulus.

In both cases the solution is constructed from the solution of a system of equations. This system is then solved by the sequential secant method, which seems to outperform comparable methods in this case. The computational procedure is then described with the particularities of this problem, e.g. how to "teach" the concept of a Riemann surface to a computer. At the end the solutions of some particular examples are displayed graphically.

Introduction. Many extremum problems have been attempted in theory and conditions for the solutions have been found but only very few solutions have been explicitly calculated. That is, however, often quite possible in our time of modern computers as the present paper shall show. We chose two particular problems partly because they have applications outside mathematics and partly because the pure mathematics involved is not too complicated. They can be considered typical though because most extremal conformal mappings satisfy equations similar to our equation (2).

Even specialists would be hard put to name five essentially distinct conformal mappings that have any kind of extremum property. The experience of actually seeing such mappings and the corresponding heuristic insight is therefore extremely limited. Originally our only aim was to fill this gap. While doing so we encountered several interesting problems and developed techniques which we believe are interesting for their own sake.

To cite Hamming [7], discussing the impact of computers on mathematics: "Much of mathematics has arisen from observation of special cases. Computers now enable us to compute many more special cases than we could by hand, to see much more detail in those we do examine, and consequently have led to many more insights."

Since [6] we have mainly improved our numerical techniques and we have replaced binary search by the sequential secant method. This method seems to outperform the discretized Newton method in our case by about 2 : 1 and binary search by a much larger factor. The fact that it does not necessarily converge even locally (unless special provisions are made) did not matter to us since we were not so much interested in proving that the method would work in every single case as in solving as many problems as possible. It deserves to be pointed out that this local nonconvergence never actually happened to us.

In § 1 we describe the mathematical background of our problems; in § 2 we describe the sequential secant method and compare it with the discretized Newton method, and § 3 describes programming techniques and discusses some examples.

* Received by the editors August 15, 1975, and in revised form May 18, 1976.

† Department of Mathematics, University of Calgary, Calgary, Alberta, Canada T2N 1N4. This work was supported by the National Research Council of Canada.

A reader mainly interested in pure mathematics can skip § 2 and most of § 3. Sections 2 and 3 should also be understandable on their own. They provide numerical techniques which we believe are interesting in their own right.

1. Mathematical background.

1.1. The extremum problems. Already in [6] we treated the following extremum problem: Given n points c_i in the complex plane $\mathbb{C}$, find a continuum that contains all the c_i and has minimal capacity. We showed there that the solution can be obtained from the following system of $2n-4$ real equations in the $n-2$ unknown complex numbers α_i $(i=1,\cdots,n-2)$:

$$\text{(1)}\qquad \begin{aligned} &\operatorname{Re}\int_{a_1}^{a_i}\sqrt{Q(z)}\,dz=0, \qquad i=2,\cdots,n-2,\\ &\operatorname{Re}\int_{a_1}^{c_i}\sqrt{Q(z)}\,dz=0, \qquad i=1,\cdots,n-1,\end{aligned}$$

where

$$Q(z)=\prod_{j=1}^{n-2}(z-a_j)\Big/\prod_{j=1}^{n}(z-c_j).$$

We used the method of binary search but it converged so slowly that we could only solve a limited number of cases. The sequential secant method converges much faster and gives a drastically greater scope. We shall refer to this problem as *Problem* 1.

In this paper we will mainly deal with a conformal invariant of a doubly connected domain bounded by two continua C and D. Such a domain can be mapped conformally onto an annulus $\{1<|\zeta|<r\}$ where the number r is uniquely determined by the domain. (See [1, pp. 246–247].)

$\log r$ is called the modulus of the domain and arises in many problems of complex variables. It is also the reciprocal of the capacity of the two-dimensional capacitor determined by the two complementary continua. It is connected with the capacity of one continuum in the following way: If we leave C fixed and let $D=\{|z|\geqq R\}$ then $\operatorname{cap} C=\lim_{R\to\infty}[\log r-\log R]$ (see [8], [10]).

We now pose the following extremum problem: Given two clusters of points $\{c_1,\cdots,c_n\}$ and $\{d_1,\cdots,d_m\}$, find continua C and D which contain all the c_i (respectively d_i) such that the mutual capacity is minimal. We shall refer to this problem as *Problem* 2. It should be pointed out that ∞ does not play the distinguished role it played in Problem 1. It is therefore more convenient to use the extended plane $\bar{\mathbb{C}}$.

There is always a solution to this problem but in general more than one. In fact if the two clusters are $\{-1,+1\}$ and $\{0,\infty\}$ then it is known that one solution is the upper half of the unit circle-line plus the lower half of the imaginary axis, and another solution is the lower half of the unit circle-line plus the upper half of the imaginary axis. One observes that the resulting two doubly connected domains cannot be deformed into each other within the four times punctured Riemann sphere $\bar{\mathbb{C}}\setminus\{+1,-1,0,\infty\}$. This observation gives rise to the following definition:

DEFINITION. Two regions Ω_1 and Ω_2 in $\bar{\mathbb{C}}\backslash\{c_i, d_i\}$ are called *homotopic* in $\bar{\mathbb{C}}\backslash\{c_i, d_i\}$ if there is a continuous mapping $h: \Omega_1 \times [0, 1] \to \bar{\mathbb{C}}\backslash\{c_i, d_i\}$ such that $h(z, 0) = z$ and $h(z, 1)$ is a homeomorphism of Ω_1 onto Ω_2.

It is easy to show that this is an equivalence-relation. One can therefore speak of homotopy-classes. We restate our problem in the following manner:

Problem 2: Given $\{c_i\}$ and $\{d_i\}$ and a homotopy class H that separates the $\{c_i\}$ from the $\{d_i\}$, find continua C and D containing the c_i (respectively d_i) such that $\Omega = \bar{\mathbb{C}}\backslash(C \cup D) \in H$ and that the capacity of Ω is minimal.

It is known [5], [13] that this problem always has a unique solution.

1.2. The fundamental equations. We start with the known fact (see [5], [12]) that the mapping function which maps Ω onto $\{1 < \zeta| < r\}$ satisfies an equation of the form

$$\log \zeta = \int \sqrt{\frac{P(x)}{\prod (z - c_i) \prod (z - d_i)}}\, dz \tag{2}$$

where $P(z)$ is a polynomial of degree $n + m - 4$. It has $n - 2$ zeros on C and $m - 2$ zeros on D (counting multiplicities). C and D consists of analytic arcs and their limiting endpoints.

Those arcs are the trajectories of $Q(z)\, dz^2 \geqq 0$, where $Q(z)$ is the rational expression under the square root. The endpoints are the poles of Q where precisely one arc ends. They are among the (but in singular cases not necessarily all) c_i and d_i.

On a zero order k precisely $k + 2$ arcs meet at equally spaced angles (see [8]). We shall refer to such a point as branch point of multiplicity k. These conditions are also sufficient for extremality. Furthermore C and D do not contain any closed curves. For proofs of these statements see [5], [13].

We denote the zeros of P on C with a_i and the zeros of P on D with b_i. Since Re $(\log \zeta) = 0$ on C and Re $(\log \zeta) = \log r = \text{const.}$ on D it follows that there are polygonal arcs δ_j joining a_1 with c_j for $1 \leqq j \leqq n$ and a_1 with a_k $(k = j - n + 1)$ for $n + 1 \leqq j \leqq 2n - 3$ and polygonal arcs γ_j joining b_1 with d_j for $1 \leqq j \leqq m$ and b_1 with b_k $(k = j - m + 1)$ for $m + 1 \leqq j \leqq 2m - 3$ and a real number α representing the argument of the leading term of $P(z)$ such that

$$\tag{3} \begin{aligned} &\operatorname{Re} \int_{\delta_j} \sqrt{Q(z)}\, dz = 0, \qquad j = 1, \cdots, 2n - 3,\\ &\operatorname{Re} \int_{\gamma_j} \sqrt{Q(z)}\, dz = 0, \qquad j = 1, \cdots, 2m - 3, \quad j \neq m, \end{aligned}$$

where

$$Q(z) = e^{i\alpha} \frac{\prod (z - a_i) \prod (z - b_i)}{\prod (z - c_i) \prod (z - d_i)}$$

and where the branch of the square root is followed continuously along the arcs.

It is to be observed that we dropped an equation corresponding to a path that connects b_1 with d_m, thus getting $2(n + m) - 7$ real equations in as many real unknowns. We shall show that this last equation is automatically satisfied.

It should also be pointed out that the δ_j and γ_j can be chosen such that $\bar{\mathbb{C}}\backslash\{\bigcup\delta_j, \bigcup\gamma_j\}\in H$, or more explicitly that H is determined by these polygons.

1.3. Sufficiency of the fundamental equations. In this subsection we shall show the following:

THEOREM 1. *Given a real number α and complex numbers a_i and b_i and polygonal arcs δ_j and γ_j as at the end of the last subsection such that no δ_j intersects a γ_j and such that equations* (3) *are satisfied. Then the missing equation*

$$\operatorname{Re}\int_{\gamma_m}\sqrt{Q(z)}\,dz=0$$

holds and the a_i and the b_i are the branch points of the extremal continua in the homotopy class of $\bar{\mathbb{C}}\backslash\{\bigcup\gamma_j, \bigcup\delta_j\}$.

The proof follows essentially the lines of Theorem 1 of [6] but it is more involved because the topology is more complicated.

It involves the concepts of a covering surface which is extensively discussed in any text on Riemann surfaces.

We start the proof by constructing a simply connected covering surface of the Riemann sphere $\bar{\mathbb{C}}$ on which $|\operatorname{Re}\int_{a_1}^{p}\sqrt{Q(z)}\,dz|$ is unique. We denote by R_1 the universal covering surface of $\bar{\mathbb{C}}\backslash\{d_i, b_i\}$. R_1 contains for each c_i and a_i infinitely many $\hat{c}_{ij}$ and $\hat{a}_{ij}$ covering them. We choose one such copy for $a_1:\hat{a}_1$. Then we lift all the δ_i with initial point $\hat{a}_1$, this obtaining unique endpoints $\hat{c}_i$ and $\hat{a}_i$. We now puncture R_1 on all the remaining $\hat{c}_{ij}$ and $\hat{a}_{ij}$ and denote by $\hat{R}$ the universal covering surface of this punctured R_1. $\hat{R}$ covers also the Riemann sphere and contains to each c_i and a_i precisely one point $\hat{a}_i$ and $\hat{c}_i$ covering it and those are connected by the $\hat{\delta}_i$.

We show now that if $\hat{\alpha}$ and $\hat{\beta}$ are two arcs connecting $\hat{a}_1$ and $\hat{p}$ on $\hat{R}$ then

$$|\operatorname{Re}\int_{\alpha}\sqrt{Q(z)}\,dz|=|\operatorname{Re}\int_{\beta}\sqrt{Q(z)}\,dz|$$

or, since we can change the sign of the square root all along $\hat{\beta}$ if necessary

$$\operatorname{Re}\int_{\hat{\alpha}\hat{\beta}^{-1}}\sqrt{Q(z)}\,dz=0.$$

According to Cauchy's integral theorem it is enough to show this for a set that generates the homotopy-group of $\hat{R}\backslash\{\hat{a}_2,\cdots,\hat{a}_{n-2},\hat{c}_1,\cdots,\hat{c}_n\}$. But the $\hat{\delta}_i$ followed by a small circle along the endpoints and back along $\hat{\delta}_i^{-1}$ do generate this homotopy-group and have according to our assumption the required property. (See also the first Lemma of [6].) Therefore $|\operatorname{Re}\int_{a_1}^{p}\sqrt{Q(z)}\,dz|=u(\hat{p})$ is indeed unique on $\hat{R}$. $u(\hat{p})$ is harmonic except on the set $\hat{C}=\{\hat{p}|u(\hat{p})=0\}$. $\hat{C}$ is a closed set, contains all the $\hat{c}_i$ and $\hat{a}_i$ and consists of arcs covering the trajectories of $Q(z)\,dz^2<0$, i.e. $\operatorname{Re}\sqrt{Q(z)}\,dz=0$ since those are precisely the lines $u(\hat{p})=\text{const}$. It contains no interior points since then $\operatorname{grad} u(\hat{p})=\pm\sqrt{Q(z)}=0$ on an open set, which is impossible. There are no closed curves contained on $\hat{C}$ since those divide $\hat{R}$ into two components, one of which would have to have compact closure. According to the maximum principle u would have to be zero on this component.

It is easy to show by induction on K that a connected set consisting of arcs and not containing any closed curves has precisely $K-2$ branch-points (counting multiplicity) if it has K endpoints. $\hat{C}$ has n endpoints and $n-2$ branch points; it can therefore contain only one component i.e., $\hat{C}$ is connected.

We can project $\hat{C}$ down to the Riemann-sphere and obtain a continuum C that consists of finitely many trajectories of $Q(z)\,dz^2 \leqq 0$. $\hat{C}$ can not cover any point of C twice since that would introduce new branch points. The only branch-points available are the a_i that are already taken and the b_i which are not even covered by $\hat{R}$. C is homotopic in $\bar{\mathbb{C}}\backslash\{d_i\}$ to the δ_i since $\hat{C}$ ends on the same endpoints on $\hat{R}$ and therefore on R_1 as the $\hat{\delta}_i$.

$\Omega_1 = \bar{\mathbb{C}}\backslash C$ is again simply connected and contains all the b_i and d_i. Almost as before we prove that $v(z) = |\mathrm{Re} \int_{b_1}^{z} \sqrt{Q(z)}\,dz|$ is unique in Ω_1, the only difference being that we have no equation in the system (3) that corresponds to a path joining b_1 and d_m, i.e. we need one extra path to generate the fundamental group of Ω_1. We choose a path σ joining b_1 with C, then once around on a path δ close to C and then back to b_1 along σ^{-1}. According to the argument principle, $\arg Q(z)$ increases along C by $2\pi\{(n-2)-n\} = -4\pi$ (or $2\pi\{(m-2)-m\} = -4\pi$ if $\infty \in C$) and therefore the branch of $\sqrt{Q(z)}$ remains unchanged so that integrals along σ and σ^{-1} cancel each other. Since $\mathrm{Re}\sqrt{Q(z)}\,dz = 0$ on C we get

$$\mathrm{Re} \int_{\sigma^{-1}\delta\sigma} \sqrt{Q(z)}\,dz = 0.$$

Thus we can conclude as before that $v(z)$ is unique in Ω_1. It is again harmonic except on the set $D = \{z \mid v(z) = 0\}$ plus possibly d_m. But if d_m were not in D it would be an isolated singularity of a harmonic function that is bounded in a neighborhood of d_m. This is not possible and therefore $d_m \in D$. D contains as before no closed curves and no interior points and connects all the b_i and d_i. It is harmonic in $\Omega = \bar{\mathbb{C}}\backslash\{C \cup D\}$, zero on D and constant on C. Let $\tilde{v}$ be a (multi-valued) conjugate harmonic function and X its period. Then

$$\zeta = \exp\left(\frac{2\pi}{X}\{v(z) + i\tilde{v}(zx)\}\right)$$

maps Ω conformally onto an annulus $\{1 < |z| < r\}$. This function obviously satisfies

$$\log \zeta = \int \sqrt{(2\pi/X)^2 Q(x)}\,dz$$

and therefore the extremal function according to (2).

For the discussion of this mapping we observe that the images of the concentric circles around the origin are precisely the lines $\mathrm{Re} \int \sqrt{Q(z)}\,dz = v(z) =$ const. or $Q(z)\,dz^2 < 0$. Similarly the lines $Q(z)\,dz^2 > 0$ would be the images of the radii. We added in some of our plots some of the lines of the first kind to give a better impression of the mapping. Since v is the potential function if D is charged with a charge x they are the equipotential lines. We will refer to $\mathrm{Re} \int_C^D \sqrt{Q(z)}\,dz$ as *total potential.*

Remark 1. An interesting set of questions is of course the dependence on H; especially which of the homotopy classes gives actually minimum capacity.

Also interesting is to solve for a particular homotopy class. We avoided these questions partly because we wanted to solve the easiest problem first; and that is to start with straight lines, instead of complicated polygons. It is also not quite clear whether H would switch or not during the iteration process described in § 2 if we would attempt such questions.

Remark 2. Even though Theorem 1 holds for multiple zeros without change we rather avoided this case because we suspect that the Jacobian then is zero and the iteration method becomes very inaccurate. For example, in case of Problem 1 for the square i.e. $c_i = \pm 1 \pm i$, a shift of the points according to $\pm(1+\varepsilon) \pm i$ makes the zeros shift along the real axes for $\varepsilon > 0$ and along the imaginary axis for $\varepsilon < 0$, indicative of a singular Jacobian. We solved Problem 1 numerically with the method and found that the accuracy of the solution was of the order 10^{-4} even though the function values were of the order 10^{-12}.

2. On the sequential secant method.

2.1. On problems and the known equation-solvers. We are faced with problems of solving n nonlinear equations in n unknowns where the computing costs for obtaining values of the integrals (3) for given approximations (in our cases approximations to $\operatorname{Re} a_i$, $\operatorname{Im} a_i$, $\operatorname{Re} b_i$, $\operatorname{Im} b_i$, α) are high. In addition the calculation of derivatives is very cumbersome so we had to look for methods that involve no derivatives and require only one function-value at each step.

The discretized Newton method as it is described in Ortega and Rheinboldt [11] is a possible candidate. If the discretization is as in [11, eq. (16), p. 186], then it requires $n+1$ function evaluations at each step. At the expense of speed of convergence one may skip evaluation of some of the discretized Newton method derivative matrices. It is not clear, however, that this will improve the total cost.

In the following we will discuss the sequential secant method by Wolfe [15]. This method requires only one function evaluation per iteration. Although it converges slower than the discretized Newton method, the total cost seems to be lower if it is measured by the number of function evaluations only.

For test purposes we applied the sequential secant method to Problem 1 where the assumption of the cost function is satisfied. (Each evaluation contains in our case, approximately $5000 \cdot n$ multiplications). We then compared it to the discretized Newton method. The results are given in § 2.3. A convergence analysis for the sequential secant method was given by Bittner [2].

2.2. The sequential secant method. We are given a system of n nonlinear equations in n unknowns:

$$v(x) = 0, \qquad v, x \in \mathbb{R}^n,$$

and $n+1$ approximations $x_1, \cdots, x_{n+1}$. Let $v_i = v(x_i)$, $i = 1, 2, \cdots, n+1$. We assume $v_1, \cdots, v_n$ is a complete set of vectors in $\mathbb{R}_n$ and can then write

$$v_{n+1} = \sum_{i=1}^{n} \lambda_i v_i. \tag{4}$$

We interpolate the function v linearly by $v^*(x)=B(x-b)$ where B is a linear map, such that $v^*(x_i)=v(x_i)$ for $i=1,\cdots,n+1$. In particular we have

$$v_{n+1}=B(x_{n+1}-b). \tag{5}$$

Using (4) we get

$$v_{n+1}=\sum \lambda_i v_i=\sum B(x_i\lambda_i)-\sum \lambda_i B(b)=B(\sum \lambda_i x_i-b\sum \lambda_i);$$

from this we get (together with (5)), assuming that B is nonsingular,

$$b=\frac{\sum \lambda_i x_i - x_{n+1}}{\sum \lambda_i - 1}=x_{n+1}+\frac{\sum \lambda_i x_i-\sum \lambda_i x_{n+1}}{\sum \lambda_i - 1}. \tag{6}$$

For numerical reasons we use the right-hand side of (6) since we supposedly compute a small correction at each step. Since b now is a zero of $v^*(x)$ we take it as a new approximation to a zero of $v(x)$ and replace the x's as follows:

$$\begin{aligned} x_i^{(k)}&=x_{i+1}^{(k-1)}, \quad i=1,\cdots,n,\\ x_{n+1}^{(k)}&=b^{(k)}, \end{aligned} \qquad k=1,2,3,\cdots.$$

2.3. Numerical example. To compare the sequential secant method to the discretized Newton method, we used our Problem 1 for 4 points c_i. We chose

$$c_1=1+i, \quad c_2=-1-i, \quad c_3=-1+i \quad \text{and} \quad c_4=1-i\alpha,$$

where α is a real parameter to be varied. We computed the solutions shown in Table 1, each time requiring the norm of the integrals to be less than 10^{-5}.

TABLE 1

α	The sequential secant method	Cost in function evaluations. Sequential secant method	Discretized Newton method	Cost in function evaluations. Discretized Newton method
-1	$-.0003004-i.000001258$ $.0003099+i.000008233$	11	$.001285-i.001419$ $-.001299+i.001439$	60
$-.75$	$-.1619+i.05818$ $.2210+i.07559$	10	No convergence	No convergence
$-.5$	$.3573+i.1651$ $-.2399+i.1071$	33	$-.2316+i.1129$ $.3510+i.1727$	40
$-.25$	$-.2822+i.1603$ $.4589+i.2966$	9	$-.2822+i.1603$ $.4589+i.2966$	20

These values were obtained using the same starting points for both methods.

Although one may not conclude anything from such a small sample of computed solutions, this, together with other experiments that we have made, suggests that the sequential secant method is about half as costly as the discretized Newton method for the kind of equations arising out of Problems 1 and 2.

We therefore chose the sequential secant method for the equation-solving.

2.4. Initial values. In the easier cases like Problem 1 for $n \leqq 4$ or Problem 2 for $n \leqq 3$, $m \leqq 3$, it is generally not too hard to find suitable approximations. A rough eyeball estimate and some perturbations thereof will usually be close enough to produce convergence of the sequential secant method.

The following result is an example of a typical case. The clusters are

C-points:	$-3+i2$	D-points:	$0+i3$
	$0+i$		$0-i$
	$-2-i3$		$2-i3$.

The input approximations to the sequential secant method were as in Table 2.

TABLE 2

approx. a_1	approx. b_1	approx. α
-3	$2+i$	.50
$-3.01+i$	$2.01+i1.02$	.55
$-3.04+i.24$	$2.035+i1.0245$	.4687
$-2.998+i2.41$	$2.0541+i.9987$	.4987
$-2.99987-i.002$	$2.00354+i.9874$	.556
$-3.004+i.00054$	$2.004127+i1.2415$	.55745

After 9 iterations the accuracy $\sum_{i=1}^{5} |v^i(b)| < 10^{-6}$ was reached. The solutions are listed below:

$$a_1 = -2.540 + i.7214,$$

$$b_1 = 2.111 - i1.510,$$

$$\alpha = .3491.$$

The following case is a case with 3 C-points and 4 D-points. In this case there was a problem in finding convergent initial approximations. However, the following procedure was used. An initial guess was made. The run of that guess would give a sequence of nonconvergent approximations. The approximation with the smallest norm was then chosen and a new set of input data consisting of perturbations of this set was used as initial approximation were chosen. Two or three repetitions of this procedure usually would suffice in order to find a convergent set of input data. It is clear that this procedure may also be automated. We chose not to automate it since the expense in programming effort would be quite large. Such a procedure (automated) would of course greatly enhance the success rate of the secant method. The following is a sample test run.

C-points:	$-3+i2$	D-points:	$0+i3$
	$0+i$		$0-i$
	$-2-i3$		$2-i3$
			$2+i2$.

The input approximations are listed in Table 3.

TABLE 3

approx. a_1	approx. b_1	approx. b_2	approx. α
$-3.14+i.00412$	$2.14+1.42$	$.41+i.12$	.02
$-3.0214+i.0241$	$2.014+i1.42$	$.241-i.124$	.0214
$-2.998+i.0241$	$2.0541+i.9987$	$.214+.024$	.0257
$-2.99987-i.002$	$2.00354+i.9874$	$-.2+i.2$	.412
$-3.01+i.01$	$2.01+i1.02$	$.01+i.5$	.5
$-3.04+i.024$	$2.035+i1.0245$	$-.02+i.02$	.241
-3	$2.+i$	0	.5
$-3.004+i.00054$	$2.004127+i1.2415$	$.24-i.124$	.024

These input approximations converged after a sequence of 30 steps to

$$a_1=-2.547+i.7232,$$

$$b_1=2.172+i2.128,$$

$$b_2=2.063-i1.448,$$

$$\alpha=.3306.$$

This procedure looks better than to just simply scan the space—in this case $\mathbb{R}^7$.

3. The computational procedure. A program was written that accomplishes the following tasks: 1) Solves equations (1) or (3) using the sequential secant method of § 2, 2) plots the minimal continuum and 3) plots equipotential lines between the continua.

3.1. Flags. Throughout the program a difficulty arises from the fact that the square root should be followed continuously, whereas computationally the square root on the CDC 6400 is given by a subroutine CSQRT delivering the principal value of the square root, i.e. with a discontinuity on the negative real axis. If W crosses the negative axis the direction of CSQRT(W) changes by π, i.e. one winds up on the other branch of the Riemann surface belonging to $\sqrt{W}$. In this case one must multiply by -1 to compensate for the jump. We therefore introduce flags taking on values $+1$ or -1. The test for crossing is accomplished by remembering the last value of a square root and then testing whether the new value is about π out of direction. Since we have small changes and continuous functions this is a reasonable test. An alternate test is to determine whether Re (new value/last value) <0, in which case CSQRT must have jumped between the branches.

Such flags were introduced along each path of integration since the square root should be followed continuously. Since a change of the branch all along one of the paths would change the sign of the corresponding integral and therefore destroy the iteration-procedure—after all, our sequential secant method assumes that the functions are C^1—such flags were also introduced in the solution process. One was attached to each integral between zeros and two to each integral between a zero and a pole. The initial value and its flag was passed to the next solution step.

The plotting of the continua between poles and zeros necessitated the use of a separate reusable flag and the plotting of the equipotential lines another three flags. In these cases a missing flag makes the plotter return to where it came from. It deserves to be mentioned that this did indeed happen in our earlier attempts.

In the following we will assume that such flags have been introduced without showing them explicitly.

3.2. Solution of equations (1) and (3). In order to evaluate the integrals of both problems numerically we have to rewrite the improper ones that connect a_1 with c_i (resp. b_1 with d_i) in the following manner:

$$\int_{a_1}^{c_i} \sqrt{Q(z)}\, dz = \int_{a_1}^{c_i} \frac{g(z)}{\sqrt{z-c_i}}\, dz = \int_{a_1}^{c_i} \frac{g(z)-g(c_i)}{\sqrt{z-c_i}}\, dz - 2g(c_i)\sqrt{a_1-c_i},$$

where $g(z)=\sqrt{Q(z)(z-c_i)}$, which is continuous at c_i. The integral on the right is now a proper integral. The proper integrals that occurred were evaluated by Simpson's rule over 20 subintervals. Since the path of integration in all cases was of order unity, the accuracy of the integral evaluation using the remainder term of Simpson's rule is of the order $(\frac{1}{20})^5 \approx 10^{-6}$. One might use a quadrature rule requiring less work for the same accuracy. We felt, however, that because of the rather complex nature of the program otherwise we would rather stay with a simple quadrature rule.

The results were used in the calculation of the increment of the sequential secant method. It should be noted that the program is built up of a main program (the sequential secant method) and function evaluation subroutines. The main program then is a real equation solver and the complex evaluations are delegated to subroutines and as such not known to the main part.

The accuracy of the equation solving is 10^{-5}, this being tested both against the sum of the absolute values of the functions and the sum of the absolute values of the change in function values. In view of the accuracy of Simpson's rule this is a reasonable and achievable tolerance.

3.3. Plotting *C* and *D*. The continua are now plotted starting from the zeros. Let a_i be simple zero, i.e. $Q(z)=(z-a_i)Q_i(z)$ where Q_i is analytic at a_i.

The trajectories satisfy, according to § 1.2,

$$Q(z)\, dz^2 \leqq 0$$

or

$$\arg Q(z) + 2 \arg dz = \pi$$

or

$$\arg Q_i(z) + \arg (z-a_i) + 2 \arg dz = -\pi - 2n\pi.$$

Setting $\arg (z-a_i) \cong \arg dz$ we get $\arg dz = \frac{1}{3}((2n+1)\pi - \arg Q_i(a_i))$. This gives three possible directions $2\pi/3$ radians apart. Careful consideration also has to be given to the choice of routine for computing arctan to ensure that the result is in the range 0–2π in order that the computations proceed properly. Having computed one of the directions, the other two are obtained by adding $2\pi/3$ to the argument and looping.

From the second step onwards the direction for the plot is given by $Q(z)\,dz^2<0$, which gives $\arg dz=\frac{1}{2}(-\arg Q(z)+\pi)$. It turns out that the plotting of the continuum by this method gives sufficient accuracy.

The continua terminates either on a pole or on a zero. In either case a tolerance less than the distance between any two zeros or poles in the continua is used to determine if the continum is getting close to its terminating point. If that is the case, a counter is set up that plots the exact number of segments necessary to reach the terminating point before the procedure terminates.

This provides incidently a control for the procedure since the trajectories would not hit their target if the a_i would be grossly wrong.

3.4. Plotting the equipotential lines. The total potential P between the two continua is now computed and divided into n equal parts corresponding to $n-1$ equipotential lines. In order to find the equipotential line with a given potential P_n we define an integral from a point c on C ($c\neq c_i$, $1\leqq i\leqq n$) to z as

$$f(z)=\int_c^z \sqrt{Q(z)}\,dz-P_n.$$

With $z_0=c$ we compute $z_{i+1}=z_i-f(z_i)/f'(z_i)$, $i=0,1,2,\cdots$, i.e. Newton's iteration formula.

From then on, a first attempt was made to plot the potential lines in the same manner as for the continum: i.e. direction given by

$$\arg Q(z)\,dz^2=\pi.$$

We define the points of the potential surface to be $z_0, z_1, \cdots$ with increments of uniform size along directions

$$\arg dz_i=\tfrac{1}{2}(\pi-\arg Q(z_i)),\qquad i=0,1,\cdots.$$

That is, $|dz_i|=\text{const.}$

It turns out that in this case the accumulated error is far too large and that a closed potential line was not plotted as a closed curve but as a spiral. Therefore one had to resort to predictor–corrector approximations of the following form:

$$z^*_{r+1}=z_{r-3}+4hz'_{r-1}+\frac{8h}{3}(z'_r-2z'_{r-1}+z'_{r-2}),\qquad \text{(predictor)}$$

$$z_{r+1,0}=z^*_{r+1},$$

$$z_{r+1,n+1}=z_{r-1}+2hz'_r+\frac{h}{3}(z'_{r+1,n}-2z'_r+z'_{r-1})\qquad n=0,1,2,\cdots.\qquad \text{(corrector)}$$

Using only one iteration step in the corrector step, we found that the plotting was more than accurate enough. The four initial approximations were computed by

$$z_0=z_0,$$

$$z_1=z_0+z_0,$$

$$z_2=z_0+(z'_0+z'_1)/2+z'_1,$$

$$z_3=z_1+(z'_1+z'_2)/2+z'_2,$$

this scheme being a bit more accurate than the simple stepping along the curve each time.

3.5. Specific examples. One would expect that the solutions of Problem 1 are "very straight". The plots confirm this; the curvature of the arcs is indeed very small. This is also the reason that the rough method of following the trajectories proved to be sufficient for the plots. Puzzling is the fact that the trajectory in Fig. 1 that ends at $5+i4$ has an inflection point. This seems to be a luxury in view of its being "very straight".

The input data for the configuration in Fig. 1 consisted of 5 points:

$$c_1=0+i5, \qquad c_4=-5-i3,$$
$$c_2=-5+i4, \qquad c_5=5-i5.$$
$$c_3=5+i4,$$

With suitable input approximations we had convergence in three steps to

$$a_1=.7552+i1.447, \quad a_2=.1078+i1.765, \quad a_3=-1.360+i1.272$$

with an error in the integral evaluations of the order 10^{-11}.

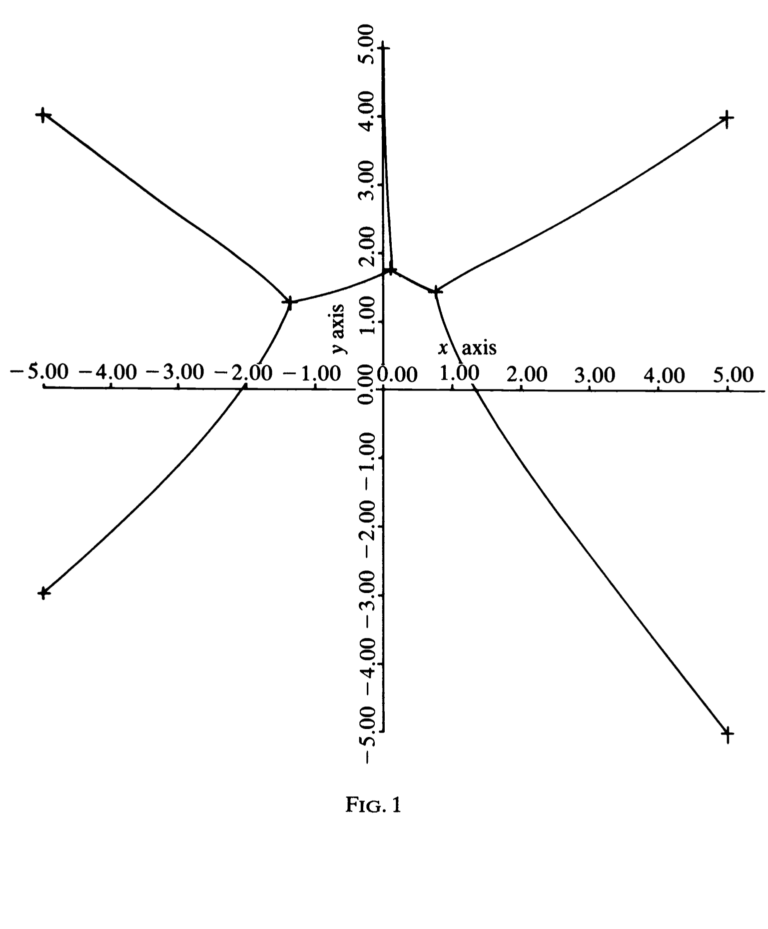

FIG. 1

An interesting behavior is shown in Figs. 2 and 3. The point c_5 is moved slightly up and the configuration changes topologically. Between the two there would be a double branch point but it seems reasonable to conjecture that in these cases the Jacobian of (1) becomes 0, and our method accordingly very inaccurate. (We can prove this fact in case of the square but not in general.)

The input data for Figs. 2 and 3 were

$$c_1=-5+i4, \qquad c_4=5-i5,$$
$$c_2=-5-i3, \qquad c_5=\begin{cases}5+i2.2 & \text{(Fig. 2)},\\ 5+i2.17 & \text{(Fig. 3)}.\end{cases}$$
$$c_3=i5,$$

The output was as shown in Table 4.

TABLE 4

	Fig. 2	Fig. 3
a_1	$-1.055+i1.504$	$-1.105+i1.448$
a_2	$2.035+i.1936$	$2.052+i1.734$
a_3	$-1.66+i1.468$	$-1.128+i1.52$

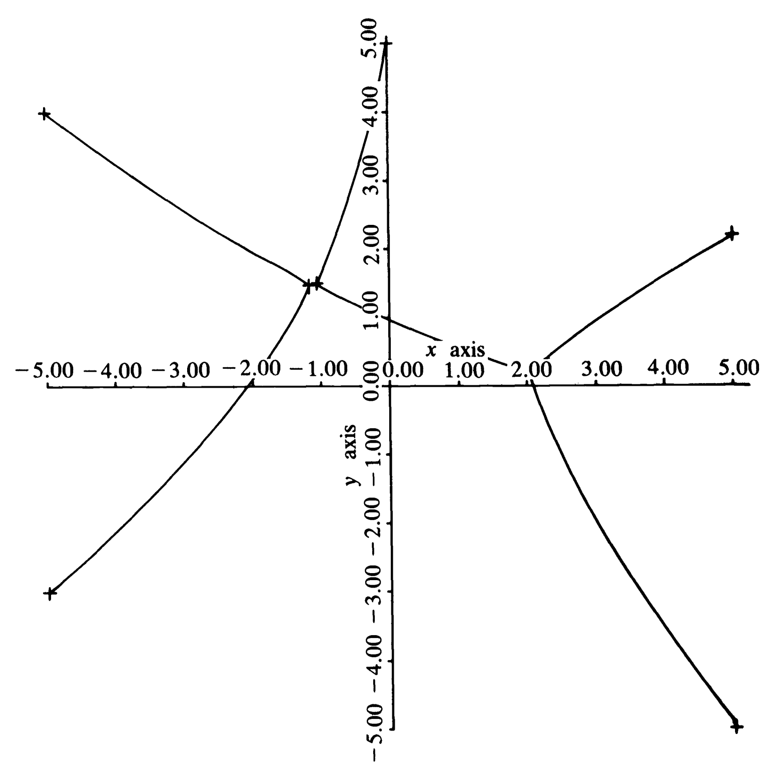

FIG. 2

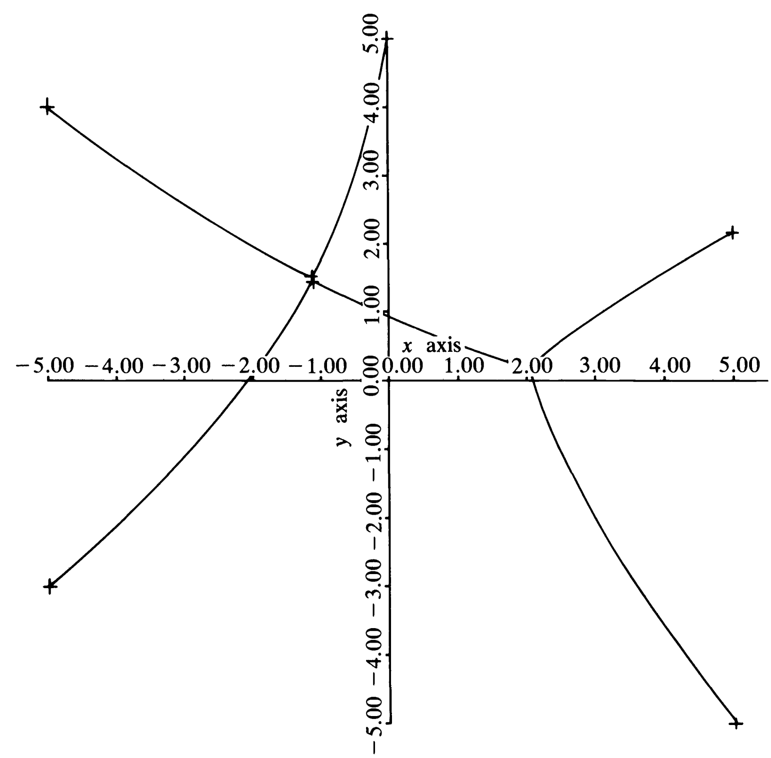

FIG. 3

Figures 4, 5, 6 and 7 display solutions to Problem 2.

As mentioned earlier, the homotopy class H is determined in these cases by the straight lines joining a_1 with a_i and c_i (respectively b_1 with b_i and d_i).

In Fig. 4 the input data was two clusters, each containing 3 points:

$$c_1 = -3 + i2, \qquad d_1 = i3,$$
$$c_3 = -1 + i, \qquad d_2 = -2,$$
$$c_3 = -2 - i3, \qquad d_3 = -2 + i.$$

With suitable approximations we got the solutions

$$a_1 = -2.984 - i.7671, \quad b_1 = -.002794 + i2.614, \quad \alpha = .9357.$$

The total potential between the two continua was calculated to be $-.6472$. Figure 4 shows the plotted continuum and the potential lines. It should be noted that

$$[\text{convex hull } C] \cap [\text{convex hull } D] \neq \varnothing.$$

The plot shows the distortion of the potential lines that one would expect in this case.

In Fig. 5 we display the case where C has three points and D 4 points:

$$\begin{aligned} c_1 &= -3+i2, & d_1 &= i3, \\ c_2 &= i, & d_2 &= -1, \\ c_3 &= -2-i3, & d_3 &= 2-i3, \\ & & d_4 &= 2+i2. \end{aligned}$$

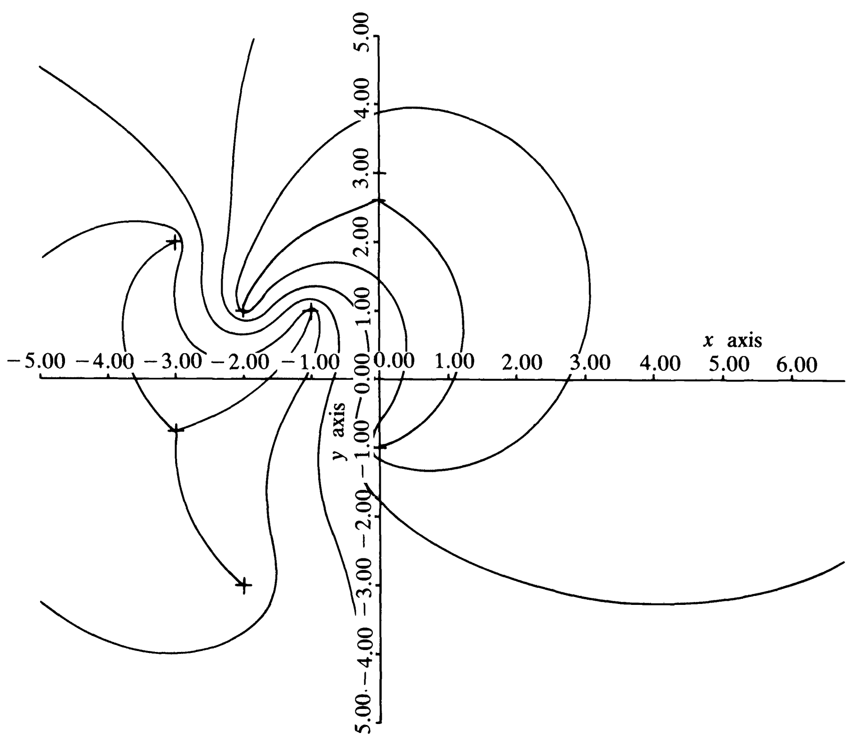

FIG. 4

Even with close input approximations it required about 30 iterations with the sequential secant method before we were close enough to the solution. This displays the fact that $\mathbb{R}^7$ is rather "roomy".

The solutions were

$$a_1 = -2.547 + i.7231, \quad b_1 = 2.173 + i2.129, \quad b_2 = 2.064 - i1.448, \quad \alpha = .3306.$$

Note that the "intrusion" of the point $+i$ into the D-continuum forces that continum into a rather unnatural shape. That is we now get a solution outside the convex hull of the D-points.

In Figs. 6 and 7 we display two solutions of the case when both continua have 4 points as follows:

$$c_1=-3, \qquad d_1=1+i3,$$
$$c_2=-1+i2, \qquad d_2=2+i,$$
$$c_3=-2-i3, \qquad d_3=-i3,$$
$$c_4=-1-i2, \qquad d_4=\begin{cases}3-i2 & \text{(Fig. 6)},\\ 2-i2 & \text{(Fig. 7)}.\end{cases}$$

The solutions were as shown in Table 5.

TABLE 5

	Fig. 6	Fig. 7
a_1	$-2.72-1.04296i$	$-2.71-i.-4485$
a_2	$-1.952-1.94i$	$-1.945-i1.941$
b_1	$2.443+1.1351i$	$2.49+i1.1173$
b_2	$2.084-2.077$	$2.897-i7.926$
α:	.006696	$-.01622$
potential:	.5740	.5629

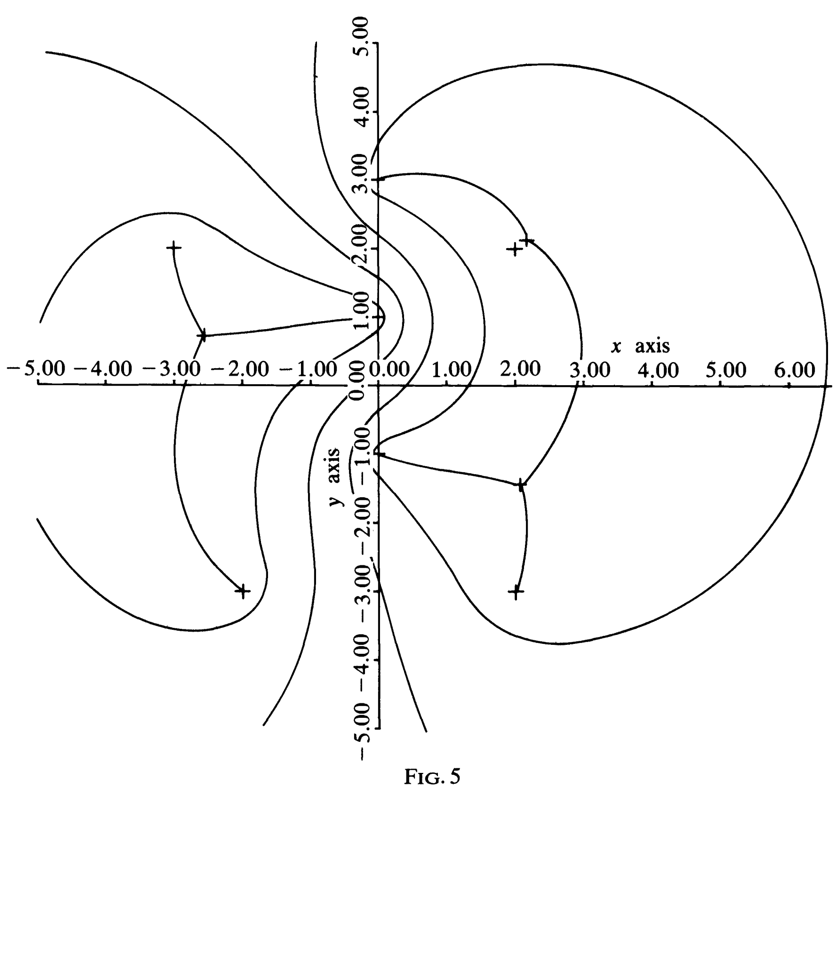

FIG. 5

It is interesting to note that even though the difference between the two sets of continua is not substantial the value of the parameter α changes surprisingly much. The parameter α has the following meaning: If β is the direction of the equipotential line at infinity, then

$$\beta = \tfrac{1}{2}[\pi - \alpha].$$

Furthermore we note that the perturbation changed one zero from being in the convex hull of D to outside the convex hull. We also notice that the continuum C did not change appreciably due to this perturbation.

In both cases the continuum C has several inflection points. This again is not surprising since D can force C into almost any shape.

The procedure normally used for the solution of the equations did not converge in the case of Fig. 7. We found that in this case, the approximation to the integral between b_1 and b_2 was not sufficiently accurate due to the proximity of a zero to the path of integration. Halving the stepsize for this particular integral solved that problem.

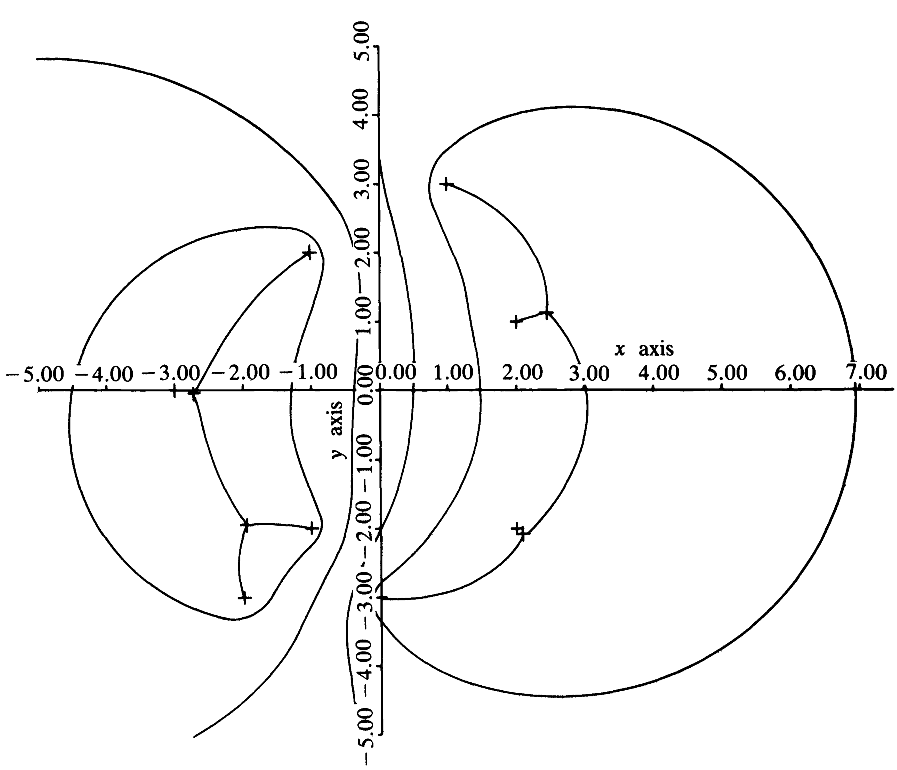

FIG. 6

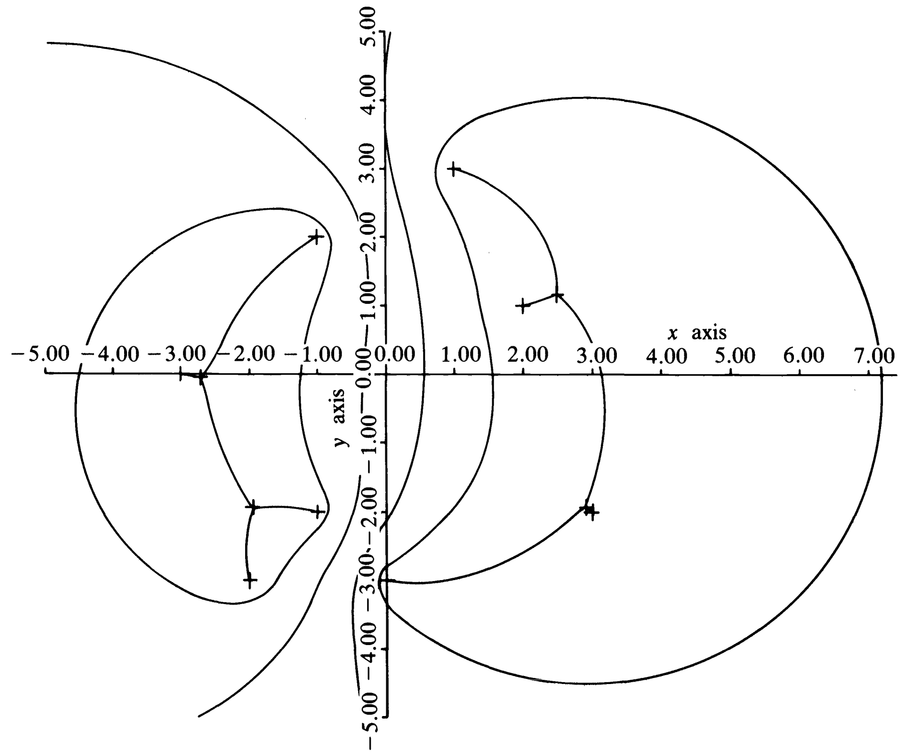

FIG. 7

REFERENCES

[1] L. V. AHLFORS, *Complex Analysis*, 2nd ed., McGraw-Hill, New York, 1966.

[2] L. BITTNER, *Eine Verallgemeinerung des Sekantenverfahrens*, Wiss., Zeitschrift, T. H. Dresden, 9 (1959), pp. 325–329.

[3] C. G. BROYDEN, *Quasi Newton or modification methods*, Numerical Solution of Nonlinear Algebraic Equations, Academic Press, New York, 1974.

[4] J. E. DENNIS, *Toward a Unified Convergence Theory for Newton-like Methods*, Nonlinear Functional Analysis and Applications, L. B. Rall, ed., Academic Press, New York, 1971.

[5] E. G. GRASSMANN, *Variationsmethoden für zweifach zusammenhängende Gebiete*, Diss. No. 4822, Eidgenössische Techn. Hoch., Zurich.

[6] E. G. GRASSMANN AND J. G. ROKNE, *An explicit calculation of some sets of minimal capacity*, this Journal, 6 (1975), pp. 242–249.

[7] R. W. HAMMING, *Impact of computers, computers and computing*, Amer. Math. Monthly, 72 (1965), no. 2 (part II).

[8] J. A. JENKINS, *Univalent Functions and Conformal mapping*, 2nd ed., Springer-Verlag, New York, 1965.

[9] ———, *On the existence of certain general extremal metrics*, Ann. of Math., 66 (1957), pp. 440–453.

[10] ———, *On certain problems of minimal capacity*, Illinois J. Math., 10 (1966), pp. 460–465.

[11] J. M. ORTEGA AND W. E. RHEINBOLDT, *Iterative Solutions of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[12] M. SCHIFFER, *On the modulus of doubly connected regions*, Quart. J. Math. Oxford Ser., 17 (1946), pp. 197–213.

[13] K. STREBEL, *Uber quadratische Differentiale mit geschlossnenen Trajectorien und extremale quasikonforme Abbildungen*, Rolf Nevanlinea Kolloquium, Festschrift, Springer-Verlag, Berlin, 1966, pp. 105–127.

[14] ———, *On quadratic differentials with closed trajectories and second order poles*, J. Analyse Math., 19 (1967), pp. 373–382.

[15] P. WOLFE, *The secant method for simultaneous non-linear equations*, Comm. ACM, 2 (1959), no. 12, pp. 12–13.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

LINEAR DIFFERENTIAL INEQUALITIES*

JAMES S. MULDOWNEY†

Abstract. A notion of generalized zero with respect to a linear differential operator L_n for a function f at a singular point of the operator was introduced by Levin and further considered by Willett. This involved a comparison of f with certain solutions of $L_n y = 0$ near the singular point. It is shown that the role of these solutions may be fulfilled by certain solutions of inequalities $L_n y \geqq 0$ ($\leqq 0$) introduced independently by Hartman and Levin. This result is applied via a generalization of the Pólya mean value theorem to the problem of finding best possible relationships between bounds on differential operators and a discussion of the extremals of these relationships. Second order operators are considered in some detail; an analogue of Landau's inequality is proved for second order operators in which the coefficients need not be constant.

1. Introduction. Real differential operators of the form

$$L_n f = f^{(n)} + a_1(t) f^{(n-1)} + \cdots + a_n(t) f, \qquad t \in (\alpha, \beta),$$

where $-\infty \leqq \alpha < \beta \leqq \infty$, are considered. It will be assumed throughout that either

(A) $f^{(n)}$ exists on (α, β) and $a_i \in C(\alpha, \beta)$, $i = 1, \cdots, n$, or

(B) $f \in \operatorname{loc} AC^{n-1}(\alpha, \beta)$ and $a_i \in \operatorname{loc} L^1(\alpha, \beta)$, $i = 1, \cdots, n$.

The results presented hold in both cases so these conditions will not usually be repeated in the statements of results and proofs. The statement $L_n f \geqq 0$ should be interpreted as holding everywhere in case (A), almost everywhere in case (B), and $L_n f \not\equiv 0$ means $L_n f(t) \neq 0$ for some $t \in (\alpha, \beta)$ in case (A) and $L_n f > 0$ (or < 0) on a set of positive measure in case (B). An end point $\tau = \beta(\alpha)$ is called *singular* if $|\tau| = \infty$ or at least one of the coefficients a_i is not integrable on a neighborhood (relative to $[\alpha, \beta]$) of τ.

Let $Z_f(I)$ denote the number of zeros of f on an interval I counting multiplicities and $Z_f(t)$ the number of zeros at t. L_n is said to be *disconjugate* on $I \subset (\alpha, \beta)$ if $Z_y(I) \leqq n-1$ for each nontrivial solution y of $L_n y = 0$. Levin [3] introduced a concept of *generalized zeros* of a function f at an end point τ which permits an extension of the notion of disconjugacy of L_n to subintervals I of $[\alpha, \beta]$ even if α and β are singular. Willett [10] also gave a definition of generalized zeros in a form convenient to the use of induction arguments for the extension of some of the classical results on disconjugate operators. It can be shown that the definitions of Levin and Willett are equivalent for functions f such that $L_n f$ does not change sign in a neighborhood of τ so they lead to the same concept of disconjugacy. Both definitions are expressed in terms of asymptotic restrictions on the behavior of f near τ involving a comparison with the behavior of members of certain solution sets (called principal systems) of $L_n y = 0$. Thus a fairly intimate knowledge of the null set of L_n is required for applications of these concepts to specific differential operators. However, results of Hartman [2] and Levin [3] show that the behavior of principal systems of solutions to $L_n y = 0$ is closely related to the behavior of principal systems of solutions to differential inequalities $L_n y \geqq 0$ ($\leqq 0$) which are more readily accessible.

* Received by the editors May 21, 1976.

† Department of Mathematics, University of Alberta, Edmonton, Alberta, Canada T6G 2G1.

The object of this paper is to present some of the ramifications of the theory of Hartman and Levin for results on singular differential operators developed by Willett [10] and later considered by Muldowney [4]. A good discussion of the results of Hartman and Levin may be found in Coppel's book [1].

2. Results. An ordered set of functions $S=(u_1, \cdots, u_n)$ is called a *Descartes system* on an interval I if $W(u_{i_1}, \cdots, u_{i_k})>0$ on I for each increasing set of indices $(i_1, \cdots, i_k)$ and $W(u_1, \cdots, u_n)$ denotes the Wronskian determinant $\det\{u_i^{(j-1)}\}$, $i, j=1, \cdots, n$. S is called a *principal system* at $\tau=\beta(\alpha)$ if $u_k>0$ in a neighborhood of τ and

$$\lim_{\tau} (u_k/u_{k+1})=0, \qquad k=1, \cdots, n-1$$

and is a *fundamental principal system* on (α, β) if $(u_1, \cdots, u_n)$ is a principal system at β and $(u_n, \cdots, u_1)$ is a principal system at α. If $(u_1, \cdots, u_n)$ is a fundamental solution set for $L_n y=0$ and a principal system at β Levin considers $Z_f(\beta)=r$ if $r=\max\{k: \lim_{\beta-} (f/u_{n-k+1})=0\}$ and, if $(u_n, \cdots, u_1)$ is a principal system at α, $Z_f(\alpha)=r$ means $r=\max\{k: \lim_{\alpha+} (f/u_k)=0\}$.

PROPOSITION 1. (a) *Suppose* $S_\alpha=(u_1, \cdots, u_n)$ *satisfies*

(i) $(-1)^{n-k}L_n u_k \geqq 0$ *on a neighborhood of* α,

(ii) $(u_n, \cdots, u_1)$ *is a principal system at* α,

(iii) $(u_1, \cdots, u_n)$ *is a Descartes system on a neighborhood of* α.

Then

$$\liminf_{\alpha+} (f/u_r) \leqq 0 \leqq \limsup_{\alpha+} (f/u_r)$$

implies $Z_f(\alpha)\geqq r$ *for any function* f *such that* $L_n f$ *does not change sign on a neighborhood of* α.

(b) *Suppose* $S_\beta=(v_1, \cdots, v_n)$ *satisfies*

(i) $(-1)^{n-k-1}L_n v_k \geqq 0$ *on a neighborhood of* β,

(ii) $(v_1, \cdots, v_n)$ *is a principal system at* β,

(iii) $(v_1, \cdots, v_n)$ *is a Descartes system on a neighborhood of* β.

Then

$$\liminf_{\beta-} (f/v_{n-r+1}) \leqq 0 \leqq \limsup_{\beta-} (f/v_{n-r+1})$$

implies $Z_f(\beta)\geqq r$ *for any function* f *such that* $L_n f$ *does not change sign on a neighborhood of* β.

If $r<n$ *in* (a) [(b)] *then* u_n $[v_1]$ *may be omitted in the hypotheses.*

Proof of (b). It is shown by Levin [3] and is easily deduced from Hartman's results [2] (cf. Coppel [1, pp. 133–134]) that under the conditions stated there exists a fundamental solution set $(U_1, \cdots, U_n)$ to $L_n y=0$ which is a principal system at β and, near β,

$$v_k=O(U_k), \qquad k=1, \cdots, n.$$

Thus, if $(f/v_k)(t_i)\to 0$ for some sequence of points $\{t_i\}$ converging to β, $(f/U_k)(t_i)\to 0$. By a result of Hartman and Levin (cf. Coppel [1, p. 128]) L_n is disconjugate on a neighborhood of β and f/U_k is monotone near β (cf. [1, p. 132]) and thus $\lim_{\beta-}(f/U_k)=0$.

Simple examples of sets S_α and S_β can be found by considering expressions of the form $L_n(e^{\lambda t})$ or more generally $L_n(t^\nu e^{\lambda t})$. Suppose for instance that $L_n(e^{\lambda t})=e^{\lambda t}P_n(\lambda, t)$ and the polynomial $P_n(\lambda, t)$ has n real zeros $\lambda_1(t), \cdots, \lambda_n(t)$ for each t and that there exist constants $\mu_0, \cdots, \mu_n$ such that $\mu_0 \leqq \lambda_1(t) \leqq \mu_1 \cdots \leqq \lambda_n(t) \leqq \mu_n$. Then $S_{-\infty}=(e^{\mu_1 t}, \cdots, e^{\mu_n t})$, $S_\infty=(e^{\mu_0 t}, \cdots, e^{\mu_{n-1} t})$ satisfy the conditions of Proposition 1 if $\mu_0<\mu_1 \cdots <\mu_n$. If $\mu_{i-1}<\mu_i=\mu_{i+1} \cdots =\mu_{i+r}<\mu_{i+r+1}$ then the functions $(e^{\mu_i t}, e^{\mu_{i+1} t}, \cdots, e^{\mu_{i+r} t})$ should be replaced by $(|t|^{r-1}e^{\mu_i t}, |t|^{r-2}e^{\mu_i t}, \cdots, e^{\mu_i t})$ in $S_{-\infty}$ and by $(e^{\mu_i t}, te^{\mu_i t}, \cdots, t^{r-1}e^{\mu_i t})$ in S_∞.

All the applications of Proposition 1 presented here use the following theorem. The reader is reminded that L_n and f satisfy either condition (A) or condition (B) of the Introduction.

THEOREM 1. *Let L_n be disconjugate on $[\alpha, \beta]$ and let $\alpha \leqq t_0 \leqq t_1 \cdots \leqq t_m \leqq \beta$. If $Z_f(t_i) \geqq r_i$ and $r_0+\cdots+r_m=k \leqq n$, then*

$$L_n f \geqq 0 \Rightarrow p_k L_{n-k} f \geqq 0$$

where $p_k(t)=\prod_{i=0}^m \sigma(t-t_i)^{r_i}$, $\sigma(t)=\operatorname{sgn} t$, $-\infty \leqq t \leqq \infty$, $L_0 y=y$, $L_{n-k}y=W(\phi_1, \cdots, \phi_{n-k}, y)/W(\phi_1, \cdots, \phi_{n-k})$ and $(\phi_1, \cdots, \phi_{n-k})$ is a basis for $\{\phi: L_n\phi=0, Z_\phi(t_i) \geqq r_i\}$. Furthermore if $L_n f \geqq 0$, then $p_k L_{n-k} f(t)>0$ at any point $t \neq t_i$ such that $L_n f \not\equiv 0$ on some interval (t_i, t) or (t, t_i) and $r_i>0$. In particular if $t_0<t_m$, $r_0, r_m>0$ and $L_n f \not\equiv 0$ on (t_0, t_m), then $p_k L_{n-k} f>0$ on $(\alpha, \beta)-\{t_0, \cdots, t_m\}$.

This theorem, a multipoint generalization of Čaplygin's inequality, is due to Pólya in the case (A) when the points t_i are nonsingular [5, Thm. V] although Pólya's formulation is somewhat different. Extensions to include singular points t_i are given by Willett [10] and Muldowney [4]; these papers consider more general boundary conditions than those given here. The conditions on $L_n f$ which imply $p_k L_{n-k} f>0$ are an improvement on those given in [4] but follow readily from a closer scrutiny of the proof in that paper. There are also restrictions placed on the asymptotic behavior of $L_{n-k} f$ near α, β and t_i by the condition $L_n f \geqq 0$, $L_n f \not\equiv 0$, but these will not be discussed in detail. The following corollaries are used in the applications.

COROLLARY 1.1. *Under the conditions of Theorem 1, if $L_n f \geqq 0$ and $L_{n-k} f(t)=0$ for some $t \neq t_i$, then $L_n f=0$ and $L_{n-k} f=0$ on every interval of the form (t_i, t) and (t, t_i) for which $r_i>0$; i.e. $f=c_1\phi_1+\cdots+c_{n-k}\phi_{n-k}$ on the union of these intervals.*

Proof. It is clear that $L_n f=0$ on every such interval since $L_n f \not\equiv 0$ on any of the intervals implies $L_{n-k} f(t) \neq 0$ by Theorem 1. Also $L_{n-k} f(t)=0$ implies the existence of constants $(c_1, \cdots, c_{n-k})$ such that, if $\tilde{f}=f-c_1\phi_1 \cdots -c_{n-k}\phi_{n-k}$, $Z_{\tilde{f}}(t) \geqq n-k$. Thus $L_n\tilde{f}=0$ and $Z_{\tilde{f}}(I) \geqq n$ where I is the union of intervals in question and, since L_n is disconjugate on I, $\tilde{f} \equiv 0$ on I.

COROLLARY 1.2. *Let L_n and t_i be as in Theorem 1. Suppose f and g are any functions such that $L_n g>0$, $Z_f(t_i) \geqq r_i$, $Z_g(t_i) \geqq r_i$, $0 \leqq i \leqq m$, $r_0+\cdots+r_m=k \leqq n$. Then*

$$p_k[L_{n-k} f+L_{n-k} g \|L_n f/(L_n g)\|] \geqq 0,$$

$$p_k[L_{n-k} f-L_{n-k} g \|L_n f/(L_n g)\|] \leqq 0,$$

and in particular,

$$|L_{n-k}f| \leqq |L_{n-k}g| \, \|L_nf/(L_ng)\|;$$

the only functions f *for which equality holds at some point are* $f = cg + c_1\phi_1 \cdots + c_{n-k}\phi_{n-k}$ *on all intervals* $[t, t_i]$, $[t_i, t]$. *Here* $\|\cdot\|$ *denotes the* $L^\infty(\alpha, \beta)$ *norm.*

Proof. Consider $\tilde{f} = f + g\|L_nf/(L_ng)\|$; then $Z_{\tilde{f}}(t_i) \geqq r_i$ and $L_n\tilde{f} = L_nf + L_ng\|L_nf/(L_ng)\| = L_ng[L_nf/(L_ng) + \|L_nf/(L_ng)\|] \geqq 0$ so that $p_kL_{n-k}\tilde{f} \geqq 0$ by Theorem 1. The same procedure with f replaced by $-f$ establishes the second inequality. The statement about achievement of equality follows from Corollary 1.1.

Remark 1. It is clear that Corollary 1.2 requires only $Z_{\tilde{f}}(t_i) \geqq r_i$ for both of the functions $\tilde{f}$ considered in the preceding proof. This is slightly less restrictive than the condition $Z_f(t_i) \geqq r_i$, $Z_g(t_i) \geqq r_i$ in the case of singular points t_i. For example if $Z_g(\beta) \geqq r$ and f satisfies the asymptotic condition of part (b) of Proposition 1, then $\tilde{f}$ also satisfies this condition so $Z_{\tilde{f}}(\beta) \geqq r$ since $L_n\tilde{f}$ does not change sign.

THEOREM 2. *Suppose* a_n *(the coefficient of* f *in* L_nf*) is either strictly positive or strictly negative on* (α, β) *and* $S_\alpha = (u_1, \cdots, u_n)$, $S_\beta = (v_1, \cdots, v_n)$ *satisfy the conditions* (i), (ii), (iii) *of Proposition* 1(a), (b) *respectively. Let* r *be an integer* $0 \leqq r \leqq n$. *In the cases* $r = 0$, $r = n$, $0 < r < n$, *it is assumed that* L_n *is disconjugate on* $(\alpha, \beta]$, $[\alpha, \beta)$, $[\alpha, \beta]$ *respectively. Suppose further that*

$$\lim_{\alpha+} u_r = \infty, \qquad \liminf_{\alpha+} \frac{f}{u_r} \leqq 0 \leqq \limsup_{\alpha+} \frac{f}{u_r}, \qquad \text{if } r > 0 \tag{2.1a}$$

and

$$\lim_{\beta-} v_{r+1} = \infty, \qquad \liminf_{\beta-} \frac{f}{v_{r+1}} \leqq 0 \leqq \limsup_{\beta-} \frac{f}{v_{r+1}}, \qquad \text{if } r < n. \tag{2.1b}$$

Then $(-1)^{n-r}a_n > 0$ *and*

$$|f| \leqq \|(1/a_n)L_nf\|. \tag{2.2}$$

The only functions for which equality holds in (2.2) *at some point are the constants. The hypotheses on* S_α *may be omitted if* $r = 0$, S_β *may be omitted when* $r = n$ *and the functions* u_n, v_1 *are required only in the cases* $r = n$, $r = 0$ *respectively.*

Proof. First the conditions imply $(-1)^{n-r}a_n > 0$. Suppose otherwise and $(-1)^{n-r}a_n < 0$; the function $f = (-1)^{n-r-1}$ satisfies (2.1) so, by Proposition 1, $Z_f(\alpha) \geqq r$, $Z_f(\beta) \geqq n - r$ and $L_nf = (-1)^{n-r-1}a_n > 0$ implies $(-1)^{n-r}f > 0$ by Theorem 1, i.e. $-1 > 0$. This contradiction shows $(-1)^{n-r}a_n > 0$. The rest of Theorem 2 follows from Corollary 1.2 and Remark 1 with $g = (-1)^{n-r}$ since $L_ng = (-1)^{n-r}a_n > 0$ and the functions

$$\tilde{f} = \pm f + g\|L_nf/(L_ng)\| = \pm f + (-1)^{n-r}\|(1/a_n)L_nf\|$$

satisfy $Z_{\tilde{f}}(\alpha) = r$, $Z_{\tilde{f}}(\beta) = n - r$.

Inequality (2.2) implies

$$\|f\| \leqq \|(1/a_n)L_nf\|. \tag{2.3}$$

The constants are not however the only extremals of (2.3). For example when

$L_n y = 0$ has a fundamental principal system of solutions $(U_1, \cdots, U_n)$ which satisfies

$$\lim_{\beta} U_i = 0, \quad i = 1, \cdots, r, \qquad \lim_{\alpha} U_i = 0, \quad i = r+1, \cdots, n$$

then, if $\gamma \in (\alpha, \beta)$, the function

$$f(t) = \begin{cases} 1 - c_1 U_1(t) - \cdots - c_r U_r(t), & t > \gamma, \\ c_{r+1} U_{r+1}(t) + \cdots + c_n U_n(t) - 1, & t < \gamma, \end{cases}$$

where the constants c_i are chosen so that $f \in \operatorname{loc} AC^{n-1}(\alpha, \beta)$, satisfies $\|f\| = 1$, $\|(1/a_n) L_n f\| = 1$.

The conditions on L_n required for Theorem 2 are satisfied by operators of the form $L_n = (D - \lambda_1) \cdots (D - \lambda_n)$ provided $\lambda_1 < \lambda_2 \cdots < \lambda_r < 0 < \lambda_{r+1} \cdots < \lambda_n$; the result also holds when $\lambda_1 \leqq \lambda_2 \cdots \leqq \lambda_r < 0 < \lambda_{r+1} \cdots \leqq \lambda_n$ although this situation is not covered by the theorem except when $r = 0$ and $r = n$. More generally if L_n satisfies the appropriate disconjugacy requirements on $[-\infty, \infty]$ and $\lambda = \lambda_i(t)$, $i = 1, \cdots, n$ are the zeros of $P_n(\lambda, t) = e^{-\lambda t} L_n(e^{\lambda t})$ then the functions $u_i(t) = e^{\mu_i t}$, $v_i(t) = e^{\nu_i t}$ satisfy the conditions of Theorem 2 provided

$$\lambda_1(t) \leqq \mu_1 \leqq \lambda_2(t) \cdots \leqq \mu_{n-1} \leqq \lambda_n(t) \leqq \mu_n, \quad \text{near } -\infty,$$

$$\nu_1 \leqq \lambda_1(t) \leqq \nu_2 \cdots \leqq \lambda_{n-1}(t) \leqq \nu_n \leqq \lambda_n(t), \quad \text{near } \infty,$$

$$\mu_i \neq \mu_j, \qquad \nu_i \neq \nu_j \quad \text{when } i \neq j, \quad \text{and} \quad \mu_r < 0 < \nu_{r+1}.$$

The exponentials should be multiplied by appropriate polynomials if $\mu_i = \mu_j$ or $\nu_i = \nu_j$. The constants μ_n, ν_i are not required except in the cases $r = n$, $r = 0$ respectively. Disconjugacy criteria may be found in the papers of Hartman [2], Levin [3] and Willett [10] and in the monograph of Coppel [1] and its references.

While inequality (2.2) is best possible in the sense that equality may hold, it can be qualitatively improved by using Theorem 1 for general $k \leqq n$ rather than for $k = n$ as was done in Theorem 2. One can replace the estimate on f in (2.2) by estimates on $M_k f$ where M_k is any operator of order $k \leqq n$. These estimates may be presented in terms of Wronskians of solutions of $L_n y = 0$ and expressions $L_j f$ where L_j are factors of L_n. However to reduce this to fairly concrete terms and still aspire to achieve sharp inequalities requires good estimates on these Wronskians. Such estimates are currently available in generality only for Wronskians of orders 1 and 2 from the work of Hartman and Levin. Thus a relatively complete picture can be presented in this paper only for $n = 2$. In situations such as $L_n = (D - \lambda_1) \cdots (D - \lambda_n)$ where $\lambda_1, \cdots, \lambda_n$ are real constants a comprehensive system of inequalities may be developed based on Theorem 1; however, good results are already available in this case in the work of Sharma and Tzimbalario [8].

Let $L_2 f = f'' + a_1(t) f' + a_2(t) f$. While the following theorems may be formulated in more general terms of systems S_α, S_β on any interval (α, β) it is more appropriate for reasons of exposition to consider only the interval $(-\infty, \infty)$ and

$$L_2(e^{\lambda t}) = \{\lambda - \lambda_1(t)\}\{\lambda - \lambda_2(t)\}\, e^{\lambda t} \tag{2.4}$$

where $\lambda_1(t)$ and $\lambda_2(t)$ are real.

PROPOSITION 2. *Suppose L_2 is of the form* (2.4) *and*

$$\nu_1 \leqq \lambda_1(t) \leqq \mu_1 < 0 < \nu_2 \leqq \lambda_2(t) \leqq \mu_2.$$

If a_i are continuous [*locally Lebesgue integrable*] *and f'' exists* [*f' is locally absolutely continuous*] *on* $(-\infty, \infty)$ *with*

$$\liminf_{-\infty} e^{-\mu_1 t} f(t) \leqq 0 \leqq \limsup_{-\infty} e^{-\mu_1 t} f(t)$$

(2.5)

$$\liminf_{\infty} e^{-\nu_2 t} f(t) \leqq 0 \leqq \limsup_{\infty} e^{-\nu_2 t} f(t),$$

then there exist continuous real valued functions ρ_1, ρ_2 on $(-\infty, \infty)$ such that

$$\nu_1 \leqq \rho_1(t) \leqq \mu_1 < 0 < \nu_2 \leqq \rho_2(t) \leqq \mu_2$$

and if b and c are any real valued functions satisfying

$$1/\rho_1 \leqq b \leqq 1/\rho_2, \qquad \rho_1 \leqq c \leqq \rho_2,$$

then

$$|bf' - f| \leqq \|(1/a_2)L_2 f\|, \tag{2.6}$$

$$|f' - cf| \leqq \frac{(\rho_2 - c)|\rho_1| + (c - \rho_1)\rho_2}{\rho_2 - \rho_1} \left\| \frac{1}{a_2} L_2 f \right\|. \tag{2.7}$$

Equality can hold only for certain functions. The constants ν_1, μ_2 may be omitted in the hypothesis and conclusion.

This result is proved by Sharma and Tzimbalario [8, Cor. 4] in the case of constant coefficient operators. Although it is best possible in that equality can hold, a qualitative improvement is given in (2.10), (2.11) and Theorem 3.

Remark 2. If $\lambda_1(t) = -\lambda_2(t)$(i.e. $a_1(t) = 0$), then it is only necessary to assume that $a_2(t) < 0$ provided (2.5) is relaced by the conditions

$$\liminf_{\pm\infty} \frac{1}{t} f(t) \leqq 0 \leqq \limsup_{\pm\infty} \frac{1}{t} f(t) \tag{2.5$'$}$$

and $\mu_2 = -\nu_1 = \|a_2\|^{1/2}$.

Proof of Proposition 2. The operator L_2 is disconjugate on $(-\infty, \infty]$ and on $[-\infty, \infty)$. Thus two positive solutions U_1, U_2 of $L_2 y = 0$, unique to within positive multiples, are determined by the conditions $Z_{U_1}(\infty) = 1$, $Z_{U_2}(-\infty) = 1$. Let $u_i(t) = e^{\mu_i t}$, $v_i(t) = e^{\nu_i t}$, $i = 1, 2$; then $(-1)^{2-i} L_2 u_i \geqq 0$, $(-1)^{1-i} L_2 v_i \geqq 0$ and (u_1, u_2) is an $S_{-\infty}$ system while (v_1, v_2) is an S_∞ system. Since $\lim_{-\infty} (v_2/u_1) = \lim_{-\infty} (u_2/u_1) = 0$, $Z_{v_2}(-\infty) \geqq 1$, $Z_{u_2}(-\infty) \geqq 1$ and, from Theorem 1, $L_2 v_2 \leqq 0$, $L_2 u_2 \geqq 0$ imply $W(U_2, v_2) \leqq 0$, $W(U_2, u_2) \geqq 0$. Similarly $W(U_1, v_1) \leqq 0$, $W(U_1, u_1) \geqq 0$. Thus

$$\nu_1 \leqq \frac{U_1'}{U_1} \leqq \mu_1 < 0 < \nu_2 \leqq \frac{U_2'}{U_2} \leqq \mu_2.$$

If $\tilde{f} = f \pm \|(1/a_2)L_2 f\|$. Then $Z_{\tilde{f}}(\pm\infty) \geqq 1$ by Proposition 1, and therefore by Corollary 2,

$$|W(U_1, f)| \leqq |W(U_1, 1)| \left\| \frac{1}{a_2} L_2 f \right\|, \qquad |W(U_2, f)| \leqq |W(U_2, 1)| \left\| \frac{1}{a_2} L_2 f \right\|.$$

Thus, if $\rho_1 = U_1'/U_1$, $\rho_2 = U_2'/U_2$,

$$\left|\frac{1}{\rho_1}f'-f\right| \leqq \left\|\frac{1}{a_2}L_2 f\right\|, \qquad \left|\frac{1}{\rho_2}f'-f\right| \leqq \left\|\frac{1}{a_2}L_2 f\right\|. \tag{2.8}$$

If $\gamma\colon (-\infty, \infty) \to [0, 1]$ and $b = \gamma(1/\rho_1) + (1-\gamma)(1/\rho_2)$ then (2.6) follows from (2.8) and the triangle inequality. Inequality (2.7) follows from

$$|f'-\rho_1 f| \leqq |\rho_1| \left\|\frac{1}{a_2}L_2 f\right\|, \qquad |f'-\rho_2 f| \leqq \rho_2 \left\|\frac{1}{a_2}L_2 f\right\|$$

similarly.

To investigate the nature of the extremals of (2.6) observe that $|(1/\rho_1)f' - f|(t_0) = \|(1/a_2)L_2 f\|$ implies $f = \eta + c_1 U_1$ on (t_0, ∞) where c_1 and η are constants, by Corollary 1.2. Similarly $|(1/\rho_2)f' - f|(t_0) = \|(1/a_2)L_2 f\|$ implies $f = \delta + c_2 U_2$ on $(-\infty, t_0)$. If both inequalities hold simultaneously then $\delta = \pm\eta$ since $|\delta| = |\eta| = \|(1/a_2)L_2 f\|$, and the constants c_1, c_2 must be chosen so that $f \in C^1(-\infty, \infty)$. When $1/[\rho_1(t_0)] < b(t_0) < 1/[\rho_2(t_0)]$, $|bf' - f|(t_0) = \|(1/a_2)L_2 f\|$ can only hold if it holds in both of the extreme cases $b(t_0) = 1/[\rho_1(t_0)]$, $b(t_0) = 1/[\rho_2(t_0)]$ and also $(bf' - f)(t_0)$ must be of the same sign in these two cases. This rules out the possibility $\delta = -\eta \neq 0$ and the only functions f for which

$$|bf'-f|(t_0) = \left\|\frac{1}{a_2}L_2 f\right\|, \qquad \frac{1}{\rho_1(t_0)} < b(t_0) < \frac{1}{\rho_2(t_0)}$$

are the constants. In (2.7) the same possibilities for equality as in (2.6) exist in the extreme cases $c(t_0) = \rho_1(t_0)$, $c(t_0) = \rho_2(t_0)$ and when $\rho_1(t_0) < c(t_0) < \rho_2(t_0)$ equality is only achieved for $f = \eta + c_1 U_1$ on (t_0, ∞), $f = -\eta + c_2 U_2$ on $(-\infty, t_0)$ with the appropriate choice of c_1, c_2.

To obtain an improvement on (2.6) consider

$$\phi_1(\eta, t) = [U_1(t)U_2'(\eta) - U_2(t)U_1'(\eta)]/[W(U_1, U_2)(\eta)],$$

$$\phi_2(\eta, t) = [U_2(t)U_1(\eta) - U_1(t)U_2(\eta)]/[W(U_1, U_2)(\eta)],$$

i.e. $L_2\phi_i(\eta, t) = 0$ and $\phi_i^{(j-1)}(\eta, \eta) = \delta_{ij}$, $i, j = 1, 2$. If $\phi_1(t) = \phi_1(0, t)$ and $\phi_2(t) = \phi_2(0, t)$ let

$$\tilde{f}(t) = f(t) - f(0)\phi_1(t) - f'(0)\phi_2(t) + \left\|\frac{1}{a_2}L_2 f\right\|(\phi_1(t) - 1);$$

since $Z_{\tilde{f}}(0) \geqq 2$ and $L\tilde{f} = a_2[(1/a_2)L_2 F - \|(1/a_2)L_2 f\|] \geqq 0$ (recall $a_2 < 0$) it follows that $\tilde{f}(t) \geqq 0$ and $\tilde{f}(t_0) = 0$ for $t_0 > 0$ $(t_0 < 0)$ if and only if $\tilde{f} \equiv 0$ on $[0, t_0]$ $([t_0, 0])$. The same remarks apply if f is replaced by $-f$ and therefore

$$\left|f'(0) + \frac{\phi_1(t)}{\phi_2(t)}f(0)\right| \leqq \frac{1}{|\phi_2(t)|}\left(|f(t)| - \left\|\frac{1}{a_2}L_2 f\right\|\right) + \frac{\phi_1(t)}{|\phi_2(t)|}\left\|\frac{1}{a_2}L_2 f\right\|. \tag{2.9}$$

Thus, if

$$c(0) \geqq -\lim_{-\infty}\frac{\phi_1}{\phi_2} = \frac{U_2'(0)}{U_2(0)} = \rho_2(0) > 0 \quad \text{or} \quad c(0) \geqq -\lim_{-\infty}\frac{\phi_1}{\phi_2} = \frac{U_2'(0)}{U_2(0)} = \rho_2(0) > 0$$

(2.9) implies

$$(2.10)\qquad |f'(0)-c(0)f(0)|\leqq G(c(0))$$

where $G(c(0))=G(L_2,f,c(0))$ is defined by

$$G(c(0))=\frac{1}{|\phi_2(t_0)|}\left(\|f\|-\left\|\frac{1}{a_2}L_2f\right\|\right)+|c(0)|\left\|\frac{1}{a_2}L_2f\right\|$$

and t_0 is defined by $\phi_1(t_0)/[\phi_2(t_0)]=-c(0)$. Under the conditions of Proposition 2 it follows from (2.6) with $b=0$ that $\|f\|\leqq\|(1/a_2)L_2f\|$ and therefore (2.10) is an improvement on (2.6). If $c_1\leqq\rho_1$ and $c_2\geqq\rho_2$ then

$$(2.11)\quad |f'(0)-c(0)f(0)|\leqq\frac{[(c_2(0)-c(0))G(c_1(0))+(c(0)-c_1(0))G(c_2(0))]}{[c_2(0)-c_1(0)]}$$

for all c, $c_1\leqq c\leqq c_2$, which is (2.7) in the case $c_1=\rho_1$, $c_2=\rho_2$; however other choices of c_1, c_2 lead to better inequalities than (2.7) if $\|f\|<\|(1/a_2)L_2f\|$. In fact the following construction shows that there exist pairs (c_1, c_2), $c_1\leqq\rho_1$, $c_2\geqq\rho_2$ for which equality may hold in (2.11). For each $\eta_1>0$ there exists $\eta_2<0$ such that

$$(2.12)\qquad \frac{\partial}{\partial t}\phi_1(\eta_2,0)=-\frac{\partial}{\partial t}\phi_1(\eta_1,0).$$

Let $c_1(0)=-\phi_1(\eta_1)/[\phi_2(\eta_1)]$, $c_2(0)=-\phi_1(\eta_2)/[\phi_2(\eta_2)]$ and define $f_0(t)=f_0(\eta_1,\eta_2,t)$ by

$$f_0(t)=\begin{cases}\phi_1(\eta_2,t)-\frac{1}{2}[\phi_1(\eta_1,0)+\phi_1(\eta_2,0)], & \eta_1\leqq t\leqq 0,\\ -\phi_1(\eta_1,t)+\frac{1}{2}[\phi_1(\eta_1,0)+\phi_1(\eta_2,0)], & 0\leqq t\leqq\eta_2,\end{cases}$$

$$f_0(t)=f_0(\eta_1),\quad t>\eta_1,\qquad f_0(t)=f_0(\eta_2),\quad t<\eta_2$$

so that $f\in\operatorname{loc} AC^1(-\infty,\infty)$ and

$$\|f_0\|=\tfrac{1}{2}[\phi_1(\eta_1,0)+\phi_1(\eta_2,0)]-1,\qquad \left\|\frac{1}{a_2}L_2f_0\right\|=\tfrac{1}{2}[\phi_1(\eta_1,0)+\phi_1(\eta_2,0)].$$

Since $L_2f_0(t)=(\operatorname{sgn} t)a_2(t)\|(1/a_2)L_2f_0\|$, $\eta_2\leqq t\leqq\eta_1$, it follows that

$$(\operatorname{sgn}\eta_i)\|f_0\|=f_0(\eta_i)=f_0(0)\phi_1(\eta_i)+f_0'(0)\phi_2(\eta_i)+(\operatorname{sgn}\eta_i)\left\|\frac{1}{a_2}L_2f_0\right\|(\phi_1(\eta_i)-1),$$

$i=1,2$, and so equality holds for f_0 in (2.10) with $c(0)=c_i(0)$, $i=1,2$. Also $f_0'(0)-c_1(0)f_0(0)$, $f_0'(0)-c_2(0)f_0(0)$ have the same sign so equality holds in (2.11) for f_0 with $c_1\leqq c\leqq c_2$. In the case $\eta_1=\infty$, $\eta_2=-\infty$, $c_1(0)=\rho_1(0)$, $c_2(0)=\rho_2(0)$ existence of the extremals in (2.11) has already been discussed in Proposition 2 (in this case $\|f_0\|=\|(1/a_2)L_2f_0\|$ and (2.11) is the inequality (2.7)). It remains to show that these extremals of (2.11) are unique; more precisely it will be shown that if equality holds in (2.11) for some $c(0)\in(c_1(0),c_2(0))\supset(\rho_1(0),\rho_2(0))$ where $c_i(0)=$

$-\phi_1(\eta_i)/[\phi_2(\eta_i)]$, $i=1, 2$, then η_1, η_2 satisfy (2.12) and f is a multiple of f_0 on $[\eta_1, \eta_2]$. If equality holds for some such $c(0)$ then it holds for all $c(0)\in [c_1(0), c_2(0)]$ and in particular (2.10) holds for $c=c_1, c_2$; therefore from (2.9) $|f(\eta_i)|=\|f\|$ and $f'(\eta_i)=0$. It may be assumed without loss of generality that $f(\eta_1)=\|f\|$. Let $\tilde{f}$ be as before so that $Z_{\tilde{f}}(0)\geqq 2$, $Z_{\tilde{f}}(\eta_1)\geqq 1$, $L\tilde{f}\geqq 0$ imply $\tilde{f}\equiv 0$ on $[0, \eta_1]$ from Corollary 1.1. Thus $(1/a_2)L_2f=\|(1/a_2)L_2f\|$ on $(0, \eta_1)$ and similarly $(1/a_2)L_2f=\pm\|(1/a_2)L_2f\|$ on $(\eta_2, 0)$. This together with the conditions that $|f(\eta_i)|=\|f\|$, $f'(\eta_i)=0$ and the continuity of f, f' at 0 shows that f is a constant multiple of $f_0(\eta_1, \eta_2, t)$ where η_1, η_2 satisfy (2.12).

The difficulty with (2.10), (2.11) is that the null set of L_2 is not always known so the inequalities are difficult to compute for general operators. Theorem 3 shows that the inequalities corresponding to (2.10), (2.11) for the operator $(D-\nu_1)(D-\mu_2)$ are, under the conditions of Proposition 2, also valid for the operator L_2.

THEOREM 3. *Under the conditions of Proposition* 2

$$|f'-cf|\leqq H(c) \tag{2.13}$$

if $c(t)\leqq \nu_1$ *or* $c(t)\geqq \mu_2$, *where* $H(c)=H(L_2, f, \nu_1, \mu_2, c)$ *is given by*

$$H(c)=|\nu_1-c|^{\mu_2/(\mu_2-\nu_1)}|\mu_2-c|^{-\nu_1/(\mu_2-\nu_1)}\left(\|f\|-\left\|\frac{1}{a_2}L_2f\right\|\right)+|c|\left\|\frac{1}{a_2}L_2f\right\|.$$

If $c_1(t)\leqq \nu_1$, $c_2(t)\geqq \mu_2$ *and* $c_1(t)\leqq c(t)\leqq c_2(t)$, *then*

$$|f'-cf|\leqq [(c_2-c)H(c_1)+(c-c_1)H(c_2)]/(c_2-c_1). \tag{2.14}$$

Equality can hold in (2.13) *and equality can hold in* (2.14) *whenever* $L_2=(D-\nu_1)(D-\mu_2)$ *on any interval* $[t_0+\eta_2, t_0+\eta_1]$, *where*

$$\eta_i=\frac{1}{\mu_2-\nu_1}\log\left|\frac{\mu_2-c_i}{\nu_1-c_i}\right|$$

and c_i *satisfy the constraint*

$$|\mu_2-c_1|^{\mu_2/(\mu_2-\nu_1)}|\nu_1-c_1|^{-\nu_1/(\mu_2-\nu_1)}=|\mu_2-c_2|^{\mu_2/(\mu_2-\nu_1)}|\nu_1-c_2|^{-\nu_1/(\mu_2-\nu_1)}. \tag{2.15}$$

Proof. Let

$$\phi_1(t)=\frac{1}{\mu_2-\nu_1}(\mu_2e^{\nu_1t}-\nu_1e^{\mu_2t}),\qquad \phi_2(t)=\frac{1}{\mu_2-\nu_1}(e^{\mu_2t}-e^{\nu_1t})$$

and $P_2(\lambda, t)=(\lambda-\lambda_1(t))(\lambda-\lambda_2(t))$. Consider

$$\tilde{f}(t)=f(t)-f(0)\phi_1(t)-f'(0)\phi_2(t)+\left\|\frac{1}{a_2}L_2f\right\|(\phi_1(t)-1);$$

$$L_2\tilde{f}=a_2\left[\frac{1}{a_2}L_2f-\left\|\frac{1}{a_2}L_2f\right\|\right]+L_2\phi_1\left[f(0)+f'(0)\frac{L_2\phi_2}{L_2\phi_1}+\left\|\frac{1}{a_2}L_2f\right\|\right]\geqq 0$$

since $a_2<0$,

$$L_2\phi_1=\frac{1}{\mu_2-\nu_1}[P_2(\nu_1, t)\mu_2e^{\nu_1t}-P_2(\mu_2, t)\nu_1e^{\mu_2t}]\geqq 0$$

and

$$-\frac{L_2\phi_2(t)}{L_2\phi_1(t)}=\frac{P_2(\nu_1, t)e^{\nu_1 t}-P_2(\mu_2, t)e^{\mu_2 t}}{P_2(\nu_1, t)\mu_2 e^{\nu_1 t}-P_2(\mu_2, t)\nu_1 e^{\mu_2 t}}$$

so that

$$\frac{1}{\nu_1}\leqq-\frac{L_2\phi_2}{L_2\phi_1}\leqq\frac{1}{\mu_2} \quad \text{and} \quad f(0)+f'(0)\frac{L_2\phi_2}{L_2\phi_1}+\left\|\frac{1}{a_2}L_2 f\right\|\geqq 0$$

by (2.6). The same analysis holds with f replaced by $-f$ and, since $Z_f(0)\geqq 2$ in both cases, Theorem 1 implies that (2.10), and hence (2.11), is satisfied with this choice of ϕ_1, ϕ_2. Inequality (2.10) is (2.13) in this case and (2.11) is (2.14) since if $-\phi_1/\phi_2=c$ then

$$|\phi_2|=|\nu_1-c|^{-\mu_2/(\mu_2-\nu_1)}|\mu_2-c|^{\nu_1/(\mu_2-\nu_1)}.$$

Also $\phi_1(\eta, t)=\phi_1(t-\eta)$, $\phi_2(\eta, t)=\phi_2(t-\eta)$ and the condition (2.12) is (2.15). Equality holds in (2.13) if f is a constant; there are however other extremals when L_2 is a constant coefficient operator. If $L_2=(D-\nu_1)(D-\mu_2)$ on $[t_0+\eta_2, t_0+\eta_1]$ the extremals for (2.14) are of the form $Cf_0(\eta_1, \eta_2, t_0+t)$, $t\in[t_0+\eta_2, t_0+\eta_1]$ discussed prior to the statement of Theorem 3; these are also extremals for (2.13).

For fixed values of $\|f\|$, $\|(1/a_2)L_2 f\|$, c, the best inequality from (2.14) may be obtained by minimizing the right hand side of (2.14) over all admissible pairs (c_1, c_2) such that $c\in[c_1, c_2]$. The exact minimum is difficult to determine for general (ν_1, μ_2). But if (ν_1, μ_2) is replaced by $(-\mu, \mu)$ where $\mu=\max\{-\nu_1, \mu_2\}$ then (2.14) holds with $H(c)=H(L_2, f, -\mu, \mu, c)$ and this minimization procedure leads to the following corollary, an analogue of Landau's inequality.

COROLLARY 3.1. *Under the conditions of Proposition* 2

$$|f'-cf|\leqq\mu K(f, L_2) \tag{2.16}$$

if

$$|c|\leqq\mu\|(1/a_2)L_2 f\|[K(f, L_2)]^{-1}$$

where $K(f, L_2)=\|f\|^{1/2}(2\|(1/a_2)L_2 f\|-\|f\|)^{1/2}$ *and* $\mu=\max\{-\nu_1, \mu_2\}$. *Equality can hold in* (2.16) *if* $L_2=D^2-\mu^2$.

This result was proved by Sharma and Tzimbalario [8] for the operator $L_2=D^2-\mu^2$; it was also proved in that case for $c=0$ by Schoenberg [7] and Muldowney [4]. The extremals for $c=0$ are given by Schoenberg [7] and these are also extremals for all c permitted here—the extremals found for (2.11) are those of Schoenberg in this case.

Remark 3. If $L_2=D^2+a_2(t)$ then (2.16) holds provided only that $a_2(t)<0$ and f satisfies (2.5)′, in which case $\mu=\|a_2\|^{1/2}$.

Remark 4. The present procedure also leads to a slight extension of Landau's inequality. If f is twice differentiable or if f' is locally absolutely continuous on $(-\infty, \infty)$, then

$$\tilde{f}(t)=f(t)-f(0)-tf'(0)+(t^2/2)\|f''\|$$

satisfies $\tilde{f}'' \geqq 0$ and $Z_{\tilde{f}}(0) \geqq 2$. The same is true if f is replaced by $-f$ and therefore

$$\left|f'(0)+\frac{1}{t}f(0)\right| \leqq \frac{1}{|t|}\|f\|+\frac{|t|}{2}\|f''\|, \quad \text{if } t \neq 0.$$

The right hand side has its minimum at $t = \pm\sqrt{2}\|f\|^{1/2}\|f''\|^{-1/2}$ and so

$$|f'-cf| \leqq \sqrt{2}\|f\|^{1/2}\|f''\|^{1/2}, \quad \text{if } |c| \leqq \frac{1}{\sqrt{2}}\|f''\|^{1/2}\|f\|^{-1/2}. \tag{2.17}$$

The extremals for Landau's inequality (i.e. $c = 0$; cf. Schoenberg [6]) are also extremals for (2.17).

Proposition 3 and Theorem 4 pertain to the situation when L_2 is of the form (2.4) and $\lambda_1(t)$, $\lambda_2(t)$ have the same sign.

PROPOSITION 3. *Suppose L_2 is of the form* (2.4) *with*

$$0 < \nu_1 \leqq \lambda_1(t) \leqq \nu_2 \leqq \lambda_2(t)$$

and f satisfies

$$\liminf_{\infty} e^{-\nu_1 t} f(t) \leqq 0 \leqq \limsup_{\infty} e^{-\nu_1 t} f(t). \tag{2.18}$$

Then there is a continuous real valued function ρ_1 on $(-\infty, \infty)$ *such that*

$$0 < \nu_1 \leqq \rho_1(t) \leqq \nu_2$$

and if b and c are real valued functions satisfying

$$0 \leqq b \leqq \frac{1}{\rho_1}, \qquad c < \rho_1,$$

then

$$|bf'-f| \leqq \|(1/a_2)L_2 f\|, \tag{2.19}$$

$$|f'-cf| \leqq (2\rho_1 - c)\|(1/a_2)L_2 f\|. \tag{2.20}$$

Equality can be achieved in (2.19) *but cannot be achieved in* (2.20) *if* $f \not\equiv 0$.

Proof. From the results of Hartman and Levin the operator L_2 is disconjugate on $(-\infty, \infty]$ so that a positive solution U_1 of $L_2 y = 0$, unique to within a positive multiple, is determined by the condition $Z_{U_1}(\infty) = 1$. Let $v_i(t) = e^{\nu_i t}$, $i = 1, 2$; (v_1, v_2) is an S_∞ system for L_2 and since $Z_{v_1}(\infty) \geqq 1$, $Lv_1 \geqq 0$ it follows from Theorem 1 that $W(U_1, v_1) \leqq 0$. Also $W(U_1, v_2) \geqq 0$; this follows from the fact that if U_2 is any solution such that $W(U_1, U_2) > 0$ then

$$\frac{W(U_1, U_2)}{U_1}\left[\frac{W(U_1, v_2)}{W(U_1, U_2)}\right]' = L_2 v_2 \leqq 0$$

so that if $W(U_1, v_2)(t_0) < 0$ then $W(U_1, v_2)(t) < 0$ (i.e. $v_2'/v_2 < U_1'/U_1$) for all $t \geqq t_0$, which contradicts the minimality of U_1 at ∞ since Hartman [2] shows there is a positive solution U such that $v_1'/v_1 \leqq U'/U \leqq v_2'/v_2$. Therefore, since $W(U_1, v_1) \leqq 0 \leqq W(U_1, v_2)$,

$$\nu_1 \leqq \rho_1 \leqq \nu_2, \quad \text{if } \rho_1 = U_1'/U_1.$$

If $\tilde{f}=f\pm\|(1/a_2)L_2f\|$ then $Z_{\tilde{f}}(\infty)\geqq 2$ by Proposition 1 and by Corollary 1.2

$$|f|\leqq\left\|\frac{1}{a_2}L_2f\right\|,\quad |W(U_1,f)|\leqq|W(U_1,1)|\,\|(1/a_2)L_2f\|.$$

Therefore

$$|f|\leqq\left\|\frac{1}{a_2}L_2f\right\|,\qquad \left|\frac{1}{\rho_1}f'-f\right|\leqq\left\|\frac{1}{a_2}L_2f\right\|. \tag{2.21}$$

Equality can hold in the first of these expressions if and only if f is a constant and in the second if and only if $f=\delta+\gamma U_1$ where δ and γ are constants. Inequalities (2.19), (2.20) follow from (2.21) by the triangle inequality; equality holds in (2.19), $0\leqq b<1/\rho_1$, if and only if f is a constant. To see that equality can not hold in (2.20) if $f\not\equiv 0$, observe that if $c<\rho_1$

$$|f'-cf|\leq|f'-\rho_1 f|+(\rho_1-c)|f|\leqq(2\rho_1-c)\|(1/a_2)L_2f\|$$

and that

$$f'(t_0)-c(t_0)f(t_0)=\pm(2\rho_1(t_0)-c(t_0))\|(1/a_2)L_2f\|$$

if and only if

$$f'(t_0)-\rho_1(t_0)f(t_0)=\pm\rho_1(t_0)\|(1/a_2)L_2f\|,\qquad f(t_0)=\pm\|(1/a_2)L_2f\|$$

with the sign + or − chosen consistently throughout. Since $\rho_1>0$ and from the nature of the extremals discussed above the only function for which this holds is $f(t)=0$, $t\in[t_0,\infty)$ and $\|(1/a_2)L_2f\|=0$ so that $f\equiv 0$.

A better inequality than (2.20) is proved by Sharma and Tzimbalario [8, Cor. 5] when L_2 is a constant coefficient operator. The analogue of (2.10) in this case improves both (2.20) and the result of [8], but has the same difficulties of explicit computation as (2.10). The following approximation of this analogue is also an improvement on (2.20) and [8].

THEOREM 4. *Suppose*

$$0<\nu_1\leqq\lambda_1(t)\leqq\nu_2\leqq\lambda_2(t)\leqq\mu_2,\qquad \mu_2>\nu_2,$$

and f satisfies (2.18). *Then*

$$|f'-cf|\leqq F(c) \tag{2.22}$$

where

$F(c)=c\|(1/a_2)L_2f\|$, *if* $c\geqq\nu_2$,

$$F(c)=|\mu_2-c|^{-\nu_2/(\mu_2-\nu_2)}|\nu_2-c|^{\mu_2/(\mu_2-\nu_2)}\left[\|f\|+\left\|\frac{1}{a_2}L_2f\right\|\right]+c\left\|\frac{1}{a_2}L_2f\right\|,\quad \text{if } c\leqq\nu_2.$$

Equality holds in (2.22) *for certain functions when* $c\geqq\nu_2$ *and when* $c<\nu_2$ *equality can hold only if* $L_2=(D-\nu_2)(D-\mu_2)$ *on an interval.*

Proof. Let

$$\phi_1(t)=\frac{1}{\mu_2-\nu_2}(\mu_2e^{\nu_1 t}-\nu_2e^{\mu_2 t}),\qquad \phi_2(t)=\frac{1}{\mu_2-\nu_2}(e^{\mu_2 t}-e^{\nu_2 t})$$

and consider the function

$$\tilde{f}(t)=f(t)-f(0)\phi_1(t)-f'(0)\phi_2(t)+\|(1/a_2)L_2f\|(1-\phi_1(t)).$$

$$L_2\tilde{f}=a_2\left[\frac{1}{a_2}L_2f+\left\|\frac{1}{a_2}L_2f\right\|\right]-L_2\phi_1\left[f(0)+f'(0)\frac{L_2\phi_2}{L_2\phi_1}+\left\|\frac{1}{a_2}L_2f\right\|\right]\geqq 0$$

since $a_2>0$,

$$L_2\phi_1(t)=\frac{1}{\mu_2-\nu_2}[\mu_2P_2(\nu_2,t)e^{\nu_2t}-\nu_2P_2(\mu_2,t)e^{\mu_2t}]\leqq 0,$$

$$-\frac{L_2\phi_2(t)}{L_2\phi_1(t)}=\frac{P_2(\mu_2,t)e^{\mu_2t}-P_2(\nu_2,t)e^{\nu_2t}}{\nu_2P_2(\mu_2,t)e^{\mu_2t}-\mu_2P_2(\nu_2,t)e^{\nu_2t}}$$

so that

$$0\leqq-\frac{L_2\phi_2}{L_2\phi_1}\leqq\frac{1}{\nu_2}\quad\text{and}\quad f(0)+f'(0)\frac{L_2\phi_2}{L_2\phi_1}+\left\|\frac{1}{a_2}L_2f\right\|\geqq 0$$

by Proposition 3. The same analysis is valid if f is replaced by $-f$ and since $Z_{\tilde{f}}(0)\geqq 2$ it follows as before that

$$\left|f'(0)+\frac{\phi_1(t)}{\phi_2(t)}f(0)\right|\leqq\frac{1}{|\phi_2(t)|}\left[\|f\|+\left\|\frac{1}{a_2}L_2f\right\|\right]-\frac{\phi_1(t)}{|\phi_2(t)|}\left\|\frac{1}{a_2}L_2f\right\|. \tag{2.23}$$

If $-\phi_1(t)/[\phi_2(t)]=c$ then

$$|\phi_2(t)|=|\mu_2-c|^{\nu_2/(\mu_2-\nu_2)}|\nu_2-c|^{-\mu_2/(\mu_2-\nu_2)}$$

and (2.22) follows from (2.23) and (2.19). The discussion of extremals is similar to that for preceding results.

If $\phi_i(t)$, $i=1,2$ are defined by $L_2\phi_i=0$, $\phi_i^{(j-1)}(0)=\delta_{ij}$ then (2.23) holds also—this is the analogue of (2.10) in this case. In the constant coefficient case Corollary 5 of [8] gives a similar bound to (2.22) except that $\|f\|$ is replaced by $\|(1/a_2)L_2f\|$ when $c\leqq\nu_2$. But, from (2.19) with $b=0$, $\|f\|\leqq\|(1/a_2)L_2f\|$, i.e. the bound in (2.22) does not exceed that of [8] and is an improvement when $c<\nu_2$ and $\|f\|<\|(1/a_2)L_2f\|$.

The following example illustrates how bounds may be obtained for M_2f if M_2 is any second order differential operator, given some information about f and L_2.

Example. Let $L_2=(D-\mu_1)(D-\mu_2)$ where μ_1,μ_2 are real constants. Suppose $p_1,p_2\in\operatorname{loc}L^1(-\infty,\infty)$ and $f\in\operatorname{loc}AC^1(-\infty,\infty)$ or $p_1,p_2\in C(-\infty,\infty)$ and f is twice differentiable.

(a) If $\mu_1<0<\mu_2$ and $f(t)=o(e^{\mu_1t})$ $(t\to-\infty)$, $f(t)=o(e^{\mu_2t})$ $(t\to+\infty)$, then

$$|f''+p_1(t)f'+p_2(t)f|\leqq|L_2f|+\|L_2f\|\left[\left|\frac{\mu_2^2+p_1(t)\mu_2+p_2(t)}{\mu_2(\mu_2-\mu_1)}\right|+\left|\frac{\mu_1^2+p_1(t)\mu_1+p_2(t)}{\mu_1(\mu_2-\mu_1)}\right|\right].$$

(b) If $0<\mu_1\leqq\mu_2$ and $f(t)=o(e^{\mu_1t})$ $(t\to\infty)$, then

$$|f''+p_1(t)f'+p_2(t)f|\leqq|L_2f|+\|L_2f\|\left[\left|\frac{\mu_1+\mu_2+p_1(t)}{\mu_2}\right|+\left|\frac{\mu_1^2+p_1(t)\mu_1+p_2(t)}{\mu_1\mu_2}\right|\right].$$

Proof. If L_n and M_n are any two differential operators of order n with leading coefficient 1, then $L_n = M_n$ if and only if $L_n u_i = M_n u_i$, $i = 1, \cdots, n$ for some set of functions $(u_1, \cdots, u_n)$ satisfying $W(u_1, \cdots, u_n) \neq 0$. Therefore if $M_2 f$ is a second order operator

$$M_2 f = \frac{W(u_1, u_2 f)}{W(u_1, u_2)} + \frac{W(u_1, f)}{W(u_1, u_2)} M_2 u_2 + \frac{W(u_2, f)}{W(u_2, u_1)} M_2 u_1$$

if $W(u_1, u_2) \neq 0$, and

$$M_2 f = \frac{W(u_1, u_2, f)}{W(u_1, u_2)} + \frac{W(u_1, f)}{W(u_1, u_2)} \left[M_2 u_2 - \frac{u_2}{u_1} M_2 u_1 \right] + \frac{f}{u_1} M_2 u_1$$

if $W(u_1, u_2) \neq 0$ and $u_1 \neq 0$. Choosing $M_2 f = f'' + p_1(t) f' + p_2(t) f$, $u_1(t) = e^{\mu_1 t}$, $u_2(t) = e^{\mu_2 t}$ and using Propositions 2, 3 to obtain bounds for $|W(u_1, f)|$, $|W(u_2, f)|$, $|f|$ yields the bounds given for $|M_2 f|$.

3. Concluding remarks. In [7] Schoenberg shows that if f satisfies $f(t) = o(e^{|\alpha t|})$ $(|t| \to \infty)$, $\alpha > 0$ and $L_2 f = f'' - \alpha^2 f$ then

$$\|f'\| \leqq \frac{1}{\alpha^2} \|L_2 f\|, \quad \text{if } \|f\| \geqq \frac{1}{\alpha^2} \|L_2 f\|,$$

$$\|f'\| \leqq \|f\|^{1/2} (2\|L_2 f\| - \alpha^2 \|f\|)^{1/2}, \quad \text{if } \|f\| < \frac{1}{\alpha^2} \|L_2 f\|.$$

However, as shown in the present paper and in [4] there are no real valued functions f satisfying the asymptotic condition $f(t) = o(e^{|\alpha t|})$ $(|t| \to \infty)$ and $\|f\| > (1/\alpha^2)\|L_2 f\|$ so the first inequality pertains only to those functions for which $\|f\| = (1/\alpha^2)\|L_2 f\|$ and the first and second inequalities are the same in that case. Therefore it is unnecessary to distinguish between the cases $\|f\| < (1/\alpha^2)\|L_2 f\|$, $\|f\| \geqq (1/\alpha^2)\|L_2 f\|$. Schoenberg's results are proved for complex valued functions satisfying the asymptotic condition and the techniques of the present paper are not applicable without a re-examination of the proof of Theorem 1 and its corollaries. Nevertheless it can be seen that the foregoing remarks also apply in the situation considered in [7] since such functions satisfy

$$f(t) = -\frac{1}{\alpha} \left[\int_{-\infty}^{t} e^{\alpha(s-t)} L_2 f(s)\, ds + \int_{t}^{\infty} e^{\alpha(t-s)} L_2 f(s)\, ds \right]$$

and hence $\|f\| \leqq (1/\alpha^2)\|L_2 f\|$. Several of the results proved by Sharma and Tzimbalario [8] also have an extra condition analogous to that of Schoenberg which is unnecessary.

The technique developed in this paper and in [4], while it leads to good inequalities without the necessity of constructing Green's functions, seems to have an undesirable element of guesswork in the discussion of extremals. In contrast, the method of Schoenberg and its development by others features an elegant treatment of extremals (cf. [6]), the nature of which is evident from the sign of the kernels involved.

The versions of Theorem 1 proved by Willett [10] and Muldowney [4] allow more general conditions at the endpoints α, β than those presented here. A condition $Z_f(\alpha)\geqq r$ may be replaced by

$$\limsup_{\alpha+} (-1)^{n-r}\frac{W(u_1, \cdots, u_j, f)}{W(u_1, \cdots, u_{j+1})}\geqq 0,$$

$j=0, \cdots, r-1$, if $(u_n, \cdots, u_1)$ is a principal system at α of solutions to $L_ny=0$. $Z_f(\beta)\geqq r$ may be replaced by

$$\limsup_{\beta-} (-1)^{r}\frac{W(u_n, \cdots, u_{n-j+1}, f)}{W(u_n, \cdots, u_{n-j})}\geqq 0,$$

$j=0, \cdots, r-1$, where $(u_1, \cdots, u_n)$ is a principal system of solutions at β. It is an attractive conjecture that, for example, instead of requiring $(u_1, \cdots, u_n)$ to be a principal system of solutions at β it would suffice if it were a system S_β satisfying conditions (i), (ii), (iii) of Proposition 1 (b). This conjecture is unfortunately true only for $r\leqq 2$. Consider the operator $L_3f=f'''$, $f=-t^2+\log t$ and $S_\beta=(u_1, u_2, u_3)=(1, t, t^2-\log t)$; $-f/u_3$, $-W(u_3, f)/[W(u_3, u_2)]$, $-W(u_3, u_2, f)/[W(u_3, u_2, u_1)]$ have limits 1, 0, 0 respectively at ∞. A principal system at ∞ of solutions to $L_3y=0$ is $(U_1, U_2, U_3)=(1, t, t^2)$ and the corresponding expressions have limits 1, 0, $-\infty$; the elements in this triple are not all nonnegative while those obtained from (u_1, u_2, u_3) are.

REFERENCES

[1] W. A. COPPEL, *Disconjugacy*, Lecture Notes in Mathematics, no. 220, Springer-Verlag, New York, 1971.

[2] P. HARTMAN, *Principal solutions of disconjugate n-th order linear differential equations*, Amer. J. Math., 91 (1969), pp. 306–362; *Corrigendum and addendum*, Ibid., 93 (1971), pp. 439–451.

[3] A. YU. LEVIN, *Non-oscillation of solutions of the equation* $x^{(n)}+p_1(t)x^{(n-1)}+\cdots+p_n(t)x=0$, Uspehi Mat. Nauk., 24 (1969), pp. 43–96 = Russian Math. Surveys, 24 (1969), pp. 43–100.

[4] J. S. MULDOWNEY, *On an inequality of Čaplygin and Pólya*, Proc. Roy. Irish Acad. Sect. A, 76 (1976), pp. 85–99.

[5] G. PÓLYA, *On the mean-value theorem corresponding to a given linear homogeneous differential equation*, Trans. Amer. Math. Soc., 24 (1924), pp. 312–324.

[6] I. J. SCHOENBERG, *The elementary cases of Landau's problem of inequalities between derivatives*, Amer. Math. Monthly, 80 (1973), pp. 121–158.

[7] ———, *Notes on spline functions VI. Extremum problems of the Landau-type for the differential operators* $D^2\pm 1$, MRC Tech. Summary Rep. no 143, Univ. of Wisconsin, Madison, 1974.

[8] A. SHARMA AND J. TZIMBALARIO, *Landau-type inequalities for some linear differential operators*, Illinois J. Math., 20 (1976), pp. 443–455.

[9] D. WILLETT, *Disconjugacy tests for singular linear differential equations*, this Journal, 2 (1971), pp. 536–545. *Errata*, Ibid., 3 (1972), p. 559.

[10] ———, *A generalization of Čaplygin's inequality with applications to singular boundary value problems*, Canad. J. Math., 25 (1973), pp. 1024–1039.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

INTEGRAL OPERATORS FOR FOURTH ORDER LINEAR PARABOLIC EQUATIONS*

PATRICK M. BROWN†

Abstract. An integral operator is constructed that maps analytic functions of two complex variables onto analytic solutions of fourth order linear parabolic equations of two space variables with analytic coefficients. The operator reduces to Bergman's operator for the fourth order elliptic equation when the solution is independent of the time variable. In the case of radial coefficients, the kernel functions of the operator are independent of the dimension and by a method of ascent, analytic solutions of three or more dimensions may be represented by a simple modification of the operator.

1. Introduction. Integral operators in the sense of S. Bergman [2] and I. N. Vekua [15] have been used extensively to represent and study analytic solutions of elliptic partial differential equations. The corresponding tool for parabolic equations would map analytic functions of two variables onto analytic solutions and would be valuable for investigating the analytic behavior of these solutions. Previous attempts by Bergman to construct an analogous theory for parabolic equations resulted in operators which had a complicated structure and did not yield an onto mapping [2, pp. 74–78]. Hill [11] also constructed an integral operator for parabolic equations, but again his operator had the disadvantage of constructing a very complicated kernel function. Recently Colton [4] has constructed an operator for second order parabolic equations in two variables which overcomes the difficulties of earlier attempts. Not only is the operator an onto mapping but its kernel is constructed through a straightforward method that is suitable for numerical techniques. In this paper we construct a Colton–Bergman operator which maps analytic functions of two variables to solutions of a fourth order parabolic equation in two space variables, namely

$$\Delta^2 u + au_{xx} + 2bu_{xy} + cu_{yy} + du_x + eu_y + fu + gu_t = 0, \tag{1.1}$$

where

$$\Delta \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}, \qquad u_x \equiv \frac{\partial u}{\partial x}.$$

We will assume that the coefficients are analytic in some polydisc in the space of two complex variables. This operator generalizes Bergman's operator for fourth order elliptic equations in the sense that when the solution of (1.1) is independent of t, the integral operator reduces precisely to Bergman's operator. (See [1].) In the second part of the paper we use the method of ascent developed by Gilbert (cf. [8], [9]) to construct an integral operator for a class of fourth order equations in $p+2$ independent variables. These operators can then be used to obtain a complete family of solutions of equations in two space variables.

* Received by the editors August 22, 1974, and in revised form April 20, 1976.

† Department of Mathematics, Indiana University, Bloomington, Indiana 47401. Now at the Department of Mathematics, College of Wooster, Wooster, Ohio 44691. This research was supported in part at Indiana University by the Air Force Office of Scientific Research through AF-AFOSR Grants 71-2205A and 74-2592 and The College of Wooster.

A possible application for the complete family of solutions is in the area of quasi-reversibility for non-well-posed problems (cf. [12]). In this technique a non-well-posed problem is approximated by a well-posed problem. For example, an initial boundary value problem for the backwards heat equation can be approximated by a well-posed initial boundary value problem for a fourth order parabolic equation which is then solved by a Fourier series expansion [13, p. 12]. Due to the difficulty of constructing the Fourier series for equations with variable coefficients, our integral operator provides an alternative method of solution that merits further investigation.

2. An integral operator for (1.1). We will now construct an integral operator that maps analytic functions of two complex variables to complex valued solutions of (1.1). Since it is convenient to work with complex notation we begin by considering the equation

$$U_{zzz^*z^*}+MU_{zz}+LU_{zz^*}+NU_{z^*z^*}+AU_z+BU_{z^*}+CU=DU_t \tag{2.1}$$

where $z=x+iy$ and $z^*=x-iy$. We shall abbreviate the left side of (2.1) by the notation $\mathbf{L}[U]$. Since x and y are allowed to assume complex values, z and z^* are independent complex variables and $z^*=\bar{z}$ only when x and y are real. The coefficients $M, L, \cdots, D$ are analytic functions of z and z^* defined for $|z|\leqq r_1$ and $|z^*|\leqq r_2$.

It should be noted that if $U=u+iv$, then equation (2.1) is equivalent to a system of two real fourth order equations. If C, D, and L are real valued when x and y are real valued, and $M=\bar{N}=M_1+iM_2$, $A=\bar{B}=A_1+iA_2$, then both equations of the system have the form of (1.1) where

$$\begin{gathered} a=4L+8M_1, \quad b=8M_2, \quad c=4L-8M_1, \\ d=16A_1, \quad e=16A_2, \quad f=16C, \quad g=16D. \end{gathered} \tag{2.2}$$

Motivated by Colton and Bergman we look for solutions to equation (2.1) of the form

$$U(z,z^*,t)=\frac{1}{2\pi i}\oint_{|\tau-t|=\delta}\int_{-1}^{1} E(z,z^*,\tau-t,s)f\left(\frac{z}{2}(1-s^2),\tau\right)\frac{ds\,d\tau}{(1-s^2)^{1/2}}, \tag{2.3}$$

where $f(z,t)$ is an arbitrary analytic function of two complex variables defined in a neighborhood of the origin in $\mathbb{C}^2$.

The first integral in equation (2.3) is a contour integral in the complex τ-plane in a counterclockwise direction around a circle with center t and radius δ where $0<\delta_0<\delta<\delta_1$. The second integral is over a path in the unit disc connecting the points $s=1$ and $s=-1$. The function $E(z,z^*,t,s)$ is required to be an analytic function of s for $|s|\leqq 1$, t for $\delta_0\leqq|t|\leqq\delta_1$ and (z,z^*) in some neighborhood of the origin in $\mathbb{C}^2$. By substituting (2.3) into (2.1) and integrating by parts with respect to

s, it follows that $E(z, z^*, \tau - t, s)$ must satisfy the differential equation

$$\begin{aligned} \mathbf{T}(E) \equiv & z^{-1}s^{-1}(1-s^2)[E_{zz^*z^*s} + ME_{zs} + LE_{sz^*} + \tfrac{1}{2}AE] \\ & + \tfrac{1}{4}z^{-2}s^{-2}(1-s^2)^2[E_{z^*z^*ss} + ME_{ss}] \\ & - z^{-1}s^{-1}[E_{zz^*z^*} + ME_z + \tfrac{1}{2}LE_{z^*} + \tfrac{1}{2}AE] \\ & - \tfrac{3}{4}z^{-2}s^{-3}(1-s^4)[E_{sz^*z^*} + ME_s] + \tfrac{3}{4}z^{-2}s^{-4}[E_{z^*z^*} + ME] + \mathbf{L}[E] \\ & - DE_t = 0. \end{aligned} \tag{2.4}$$

Also, the expressions

$$z^{-1}s^{-1}\mathbf{D}_1(E). \tag{2.5}$$

and

$$z^{-1}s^{-1}[\mathbf{D}_1(-2E_z - z^{-1}E - \tfrac{1}{2}z^{-1}s^{-1}E_s + \tfrac{1}{2}z^{-1}sE_s + \tfrac{1}{2}z^{-1}s^{-2}E) - AE - LE_{z^*}], \tag{2.6}$$

where $\mathbf{D}_1(H) = H_{z^*z^*} + MH$, must be continuous functions for sufficiently small values of z and z^*, $|s| \leqq 1$ and $\delta_0 \leqq |t-\tau| \leqq \delta_1$. Details relevant to the derivation of (2.4) are found in [1] and [2].

Solutions can also be represented by the integral operator

$$U(z, z^*, t) = \frac{1}{2\pi i}\oint_{|\tau - t| = \delta}\int_{-1}^{1} E(z, z^*, \tau - t, s) f\left(\frac{z^*}{2}(1-s^2), \tau\right)\frac{ds\, d\tau}{(1-s^2)^{1/2}},$$

where E satisfies an equation obtained from (2.4) by interchanging differentiation with respect to z and z^* and exchanging A with B, M with N, and vice versa. This differential equation will be denoted by

$$\mathbf{T}^*(E) = 0.$$

We now intend to construct two linearly independent solutions of (2.4), $E^{(I,1)}$ and $E^{(I,2)}$, having the form

$$E(z, z^*, t, s) = \frac{P^{(0)}(z, z^*)}{t} + \sum_{n=1}^{\infty} s^{2n}z^n P^{(2n)}(z, z^*, t) \tag{2.7}$$

and satisfying the initial conditions

$$E^{(I,1)}(z, 0, t, s) = 1/t, \qquad E^{(I,1)}_{z^*}(z, 0, t, s) = 0, \tag{2.8}$$

$$E^{(I,2)}(z, 0, t, s) = 0, \qquad E^{(I,2)}_{z^*}(z, 0, t, s) = 1/t. \tag{2.9}$$

The functions $P^{(2n)}$ are to be determined. Substituting (2.7) into (2.4) we obtain the following system of differential recursion equations for the $P^{(2n)}(z, z^*, t)$:

(2.10a) $$P^{(0)}_{z^*z^*}+MP^{(0)}=0,$$

(2.10b) $$P^{(2)}_{z^*z^*}+MP^{(2)}=-\frac{2}{t}[AP^{(0)}+LP^{(0)}_{z^*}+2P^{(0)}_{zz^*z^*}+2MP^{(0)}_z],$$

(2.10c)
$$\begin{aligned}P^{(2n+4)}_{z^*z^*}+MP^{(2n+4)}=\frac{-1}{n^2+2n+3/4}&[(2n+1)P^{(2n+2)}_{zz^*z^*}+(n+\tfrac{1}{2})\\ &\cdot(LP^{(2n+2)}_{z^*}+AP^{(2n+2)})\\ &+(2n+1)MP^{(2n+2)}_z+NP^{(2n)}_{z^*z^*}+BP^{(2n)}_{z^*}\\ &+CP^{(2n)}+AP^{(2n)}_z+P^{(2n)}_{zzz^*z^*}+MP^{(2n)}_{zz}\\ &+LP^{(2n)}_{zz^*}-DP^{(2n)}_t].\end{aligned}$$

(For $n=0$, $p^{(2n)}_t$ is replaced in (2.10c) by $-P^{(0)}/t^2$). Setting $Q^{(0)}(z, z^*)=P^{(0)}(z, z^*)$ and $Q^{(2n)}(z, z^*, t)=t^{n+1}P^{(2n)}(z, z^*, t)$, $n=1, 2, \cdots$, in (2.10) yields the recursion equations:

(2.11a) $$Q^{(0)}_{z^*z^*}+MQ^{(0)}=0,$$

(2.11b) $$Q^{(2)}_{z^*z^*}+MQ^{(2)}=-2t[AQ^{(0)}+LQ^{(0)}_{z^*}+2Q^{(0)}_{zz^*z^*}+2MQ^{(0)}_z],$$

(2.11c)
$$\begin{aligned}&Q^{(2n+4)}_{z^*z^*}+MQ^{(2n+4)}\\ &\quad=\frac{-t}{n^2+2n+3/4}[(2n+1)Q^{(2n+2)}_{zz^*z^*}+(n+\tfrac{1}{2})(LQ^{(2n+2)}_{z^*}+AQ^{(2n+2)})\\ &\qquad+(2n+1)MQ^{(2n+2)}_z+t(NQ^{(2n)}_{z^*z^*}+BQ^{(2n)}_{z^*}\\ &\qquad+CQ^{(2n)}+AQ^{(2n)}_z+Q^{(2n)}_{zzz^*z^*}+MQ^{(2n)}_{zz}+LQ^{(2n)}_{zz^*}-DQ^{(2n)}_t)\\ &\qquad+(n+1)DQ^{(2n)}].\end{aligned}$$

It is clear from (2.11) that the $Q^{(2n)}$ are uniquely determined. To prove the existence of the functions $E^{(I,1)}$ and $E^{(I,2)}$, it is necessary to show the convergence of the series (2.7). Due to the complicated nature of the recursion equations, we cannot directly apply the method of dominance to majorize the $Q^{(2n)}$ as done in previous problems (cf. [5]). For the fourth order elliptic equation, Bergman overcame this difficulty with two lemmas which have been modified due to the time variable and are stated here without proof. Details of the proofs or the properties of domination may be found in [1] or [2] respectively.

LEMMA 1 (Bergman). *Let M, M_1, R, R_1, be regular analytic functions of z and z^* for $|z|\leqq r_1$ and $|z^*|\leqq r_2$. Let $A_1(z)$ and $B_1(z)$ be regular analytic functions of z for $|z|\leqq r_1$. Let $\varphi(z, z^*)$ be a solution of the differential equation*

(2.12) $$\frac{\partial^2\varphi}{\partial z^{*2}}+M\varphi=R$$

satisfying the initial conditions

$$\varphi(z, 0)=A(z) \quad \textit{and} \quad \left.\frac{\partial\varphi}{\partial z^*}\right|_{z^*=0}=B(z)$$

and let $\varphi_1(z, z^*)$ be a solution of the equation

$$\frac{\partial^2 \varphi_1}{\partial z^{*2}} + M\varphi_1 = R_1$$

satisfying the conditions

$$\varphi_1(z, 0) = A_1(z) \quad \textit{and} \quad \left.\frac{\partial \varphi_1}{\partial z^*}\right|_{z^*=0} = B_1(z).$$

If $A(z) \ll A_1(z)$, $B(z) \ll B_1(z)$, $R(z) \ll R_1(z)$, and $M(z) \ll M_1(z)$, where "$\ll$" denotes domination, then

$$\varphi(z, z^*) \ll \varphi_1(z, z^*).$$

For the second lemma, assume $M(z, z^*)$ is regular for $|z| \leqq R_1$, $R_1 > r_1$, and $|z^*| \leqq r_2$. Let K and A, $A < 1$, be given positive numbers. Let $\{\mu_m\}$ be a sequence of nonnegative integers such that

$$(\mu_m + m - 2)(\mu_m + m - 1) > \frac{K r_2^2 R_1}{A(R_1 - r_1)} \quad \text{for } m = 0, 1, 2 \cdots,$$

and $\mu_m = 0$ for $m > m_0$ where m_0 is sufficiently large.

LEMMA 2 (Bergman). *Consider the differential equation*

$$\frac{\partial^2 \varphi}{\partial z^{*2}} + M\varphi = R(z, z^*, t) \tag{2.13}$$

where

$$M \ll \frac{k}{(1 - z/R_1)(1 - z^*/r_2)^2}$$

and

$$R \ll C(1 - z/r_1)^{-n}(1 - z^*/r_2)^{-m} \cdot (1 - t/2\delta_1)^{-p} \equiv C\{n, m, p\}, \qquad |t| < \delta_1.$$

If $\Phi(z, z^*, t)$ satisfies (2.17) *and if*

$$\Phi(z, 0, t) = \Phi_{z^*}(z, 0, t) = 0,$$

then $\Phi(z, z^*, t)$ is regular for $|z| \leqq r_1$, $|z^*| \leqq r_2$, $|t| < \delta_0$ and

$$\begin{aligned} \Phi(z, z^*, t) \ll & \frac{C r_2^2}{(\mu_m + m - 2)(\mu_m + m - 1)(1 - A)} (1 - z/r_1)^{-(n+1)} \\ & \cdot (1 - z^*/r_2)^{-(\mu_m + m - 2)} (1 - t/2\delta_1)^{-p}. \end{aligned} \tag{2.14}$$

In the following, it will be convenient to use the notation

$$(1 - z/r_1)^{-n}(1 - z^*/r_2)^{-m}(1 - t/2\delta_1)^{-p} \equiv \{n, m, p\}$$

and

$$\{n, m, 0\} \equiv \{n, m\},$$

where m, n, and p are integers. Using these lemmas we can now dominate the

$Q^{(2n)}$ and show convergence of the series (2.7). Since $Q^{(0)}$ is a regular solution of equation (2.11a) there is a constant $C^{(0)}$ such that

$$Q^{(0)}(z, z^*) \ll C^{(0)}\{1, 1\}. \tag{2.15}$$

Now, let K be the maximum of the coefficients of (2.1) for $|z| \leqq r_1$ and $|z^*| \leqq r_2$. Then the coefficients will be dominated by $K\{1, 1\}$. Since $M(z, z^*)$ is assumed to be regular for $|z| \leqq R_1$, $R_1 > r_1$, $|z^*| \leqq r_2$, there is a constant, also denoted by K, such that

$$\begin{aligned} M &\ll K(1-z/R_1)^{-1}(1-z^*/r_2)^{-1} \\ &\ll K(1-z/R_1)^{-1}(1-z^*/r_2)^{-2}. \end{aligned} \tag{2.16}$$

Let $R^{(2)}$ represent the right side of (2.11b). Then using the properties of dominance, (2.15) and (2.16), and the fact

$$t \ll 2\delta_1(1-t/(2\delta_1))^{-1},$$

we obtain

$$\begin{aligned} R^{(2)} &\equiv -2t[AQ^{(0)} + LQ^{(0)}_{z^*} + 2Q^{(0)}_{zz^*z^*} + 2MQ^{(0)}_z] \\ &\ll 4\delta_1 C^{(0)}[Kr_1^{-1}\{3, 3\} + 2r_1^{-1}r_2^{-2}\{2, 3\} + \tfrac{1}{2}Kr_2^{-1}\{2, 3\} + \tfrac{1}{2}K\{2, 2\}](1-t/(2\delta_1))^{-1} \\ &\ll 16\delta_1 C^{(0)}\left[\frac{1}{r_1 r_2^2} + \frac{K}{\tau}\right]\{3, 3, 1\}, \end{aligned}$$

where

$$\tau = \min\{r_1^m r_2^n;\ m, n = 0, 1, 2\}.$$

Thus by Lemma 2,

$$\begin{aligned} Q^{(2)} &\ll \frac{16\delta_1 C^{(0)}[1/(r_1 r_2^2) + K/\tau] r_2^2\{4, \mu_3+1, 1\}}{(\mu_3+1)(\mu_3+2)(1-A)} \\ &\equiv C^{(2)}\{4, \mu_3+1, 1\}. \end{aligned}$$

Let $\rho = \max\{\delta_1, 1\}$ and we obtain the bound

$$\frac{C^{(0)}}{r_1} = \frac{C^{(2)}(\mu_3+1)(\mu_3+2)(1-A)}{16\rho[1 + Kr_1 r_2^2/\tau]} \leqq \frac{C^{(2)}}{16}.$$

We will now show by induction that

$$Q^{(2k)} \ll C^{(2k)}\{3k+3, k+\nu(k)+3, 2k-1\} \tag{2.17}$$

and

$$\frac{C^{(2k-2)}}{r_1} \leqq \frac{1}{16} C^{(2k)}, \tag{2.18}$$

where the sequence of numbers $\nu(k)$ is defined by $\nu(0) = 0$, $\nu(1) = \mu_3$, $\nu(n) = \nu(n-1) + \mu_{\varepsilon_m}$, and $\varepsilon_n = n + 4 + \nu(n-1)$. We define

$$\sigma_n = \sum_{k=0}^{n} \mu_k \quad \text{and} \quad \sigma = \sum_{k=0}^{m_0} \mu_k.$$

Clearly $\nu(n)\leqq\nu(n+1)$ and $\nu(n)<\sigma_n\leqq\sigma$. It is easy to see that (2.17) and (2.18) hold for $k=0, 1$, and next we assume they are valid for $k=n$ and $n+1$. Let $R^{(2n+4)}$ be the right side of (2.11c). The induction hypothesis and the properties of dominance imply that

$$\begin{aligned}
R^{(2n+4)} \ll \frac{2\delta_1}{n^2+2n+3/4}\Big[&C^{(2n+2)}(2n+1)(3n+6)(n+4+\nu(n+1))\\
&\cdot(n+5+\nu(n+1))r_1^{-1}r_2^{-2}\{3n+7, n+6+\nu(n+1), 2n+2\}\\
&+2\delta_1C^{(2n)}(3n+3)(3n+4)(n+3+\nu(n))\\
&\cdot(n+4+\nu(n))(r_1r_2)^{-2}\{3n+5, n+5+\nu(n), 2n+1\}\Big]\\
+\frac{2\delta_1K}{\tau(n^2+2n+3/4)}\Big[&C^{(2n+4)}(n+4+\nu(n+1))(n+1/2)+(n+1/2)C^{(2n+2)}\\
&+(2n+1)(3n+6)C^{(2n+2)}\\
&+2\delta_1C^{(2n)}(n+3+\nu(n))(n+4+\nu(n))\\
&+2\delta_1C^{(2n)}(n+\nu(n)+3)+2\delta_1C^{(2n)}\\
&+2\delta_1(3n+3)C^{(2n)}+2\delta_1(3n+3)(3n+4)C^{(2n)}\\
&+2\delta_1(3n+3)(n+3+\nu(n))C^{(2n)}\\
&+(2n-1)C^{(2n)}+(n+1)C^{(2n)}\Big]\\
&\cdot\{3n+8, n+6+\nu(n+1), 2n+2\}
\end{aligned}$$

Using (2.18) for $k=n+1$, we find that

$$\begin{aligned}
R^{(2n+4)} \ll C^{(2n+2)}\frac{(4\rho^2)}{n^2}\Big[&(3n+6)^2(n+5+\nu(n+1))^2r_1^{-1}r_2^{-2}\\
&+\frac{2K}{\tau}(3n+6+\nu(n+1))^2\Big]\\
&\cdot\{3n+8, n+6+\nu(n+1), 2n+2\}.
\end{aligned}$$

Applying Lemma 2, we obtain

$$\begin{aligned}
Q^{(2n+4)} &\ll \frac{4\rho^2C^{(2n+2)}}{r_1(1-A)n^2(n+4+\nu(n+1)+\mu_{\varepsilon_{n+2}})(n+5+\nu(n+1)+\mu_{\varepsilon_{n+2}})}\\
&\quad\cdot\Big[(3n+6)^2(n+5+\nu(n+1))^2+\frac{2r_2^2r_1K}{\tau}(3n+6+\nu(n+1))^2\Big]\\
&\quad\cdot\{3n+9, n+4+\nu(n+1)+\mu_{\varepsilon_{n+2}}, 2n+2\}\\
&\ll \frac{4\rho^2C^{(2n+2)}}{r_1(1-A)n}\Big[9(n+5+\sigma)^2+\frac{2r_2^2r_1K}{\tau}\Big(\frac{3n-6+\sigma}{n+2}\Big)^2\Big]\\
&\quad\cdot\{3n+9, n+5+\nu(n+2), 2n+3\}.
\end{aligned}$$

Thus,

$$Q^{(2n+4)} \ll C^{(2n+4)}\{3n+9, n+5+\nu(n+2), 2n+3\}$$

and

$$\frac{C^{(2n+2)}}{r_1} = \frac{n^2(1-A)C^{(2n+4)}}{4\rho^2[9(n+5+\sigma)^2+(2r_2^2r_1K/\tau)[(3n+6+\sigma)/(n+2)]^2]}$$
$$\leqq \frac{n^2(1-A)C^{(2n+4)}}{4\rho^2[9(n+5+\sigma)^2]} < \frac{C^{(2n+4)}}{16}.$$

This completes the induction proof.

We now prove the convergence of the series (2.7). Let $|s| \leqq s_0$ where $s_0 \leqq 1$, $\delta_0 < |t| < \delta_1$ where $\delta_0 > 0$ and δ_1 is arbitrarily large, and $\rho = \max(\delta_1, 1)$. From (2.17) it is seen that the series expansion for $E^{(I,k)}$, $k = 1, 2$, is majorized by the series

$$\frac{C_0\{1, 1\}}{\delta_0} + \sum_{n=1}^{\infty} \frac{s_0^{2n}|z|^n}{\delta_0^{n+1}} C^{(2n)}\{3n+3, n+3+\sigma, 2n-1\}.$$

Since $|t| < \delta_1$ and $|1-t/(2\delta_1)|^{-2} < 4$, an application of the ratio test shows the series converges absolutely and uniformly for

$$\left|\frac{z}{r_1}\right| \left|1-\frac{z}{r_1}\right|^{-3} \left|1-\frac{z^*}{r_2}\right|^{-1} < \frac{(1-A)\delta_0}{144\rho^2 s_0^2}.$$

We summarize this result in the following lemma.

LEMMA 3. *There exist two sequences of functions,*

$$\{P^{(I,i,0)}(z, z^*), P^{(I,i,2n)}(z, z^*, t), n = 1, 2, \cdots\}, \qquad i = 1, 2$$

satisfying the differential equations (2.10 *a, b, c*) *such that*

$$P^{(I,1,0)}(z, 0) = 1, \qquad P_{z^*}^{(I,1,0)}(z, 0) = 0,$$

(2.19)
$$P^{(I,1,2n)}(z, 0, t) = 0, \qquad P_{z^*}^{(I,1,2n)}(z, 0, t) = 0, \qquad n = 1, 2, \cdots,$$

and

(2.20)
$$P^{(I,2,0)}(z, 0) = 0, \qquad P_{z^*}^{(I,2,0)}(z, 0) = 1,$$
$$P^{(I,2,2n)}(z, 0, t) = 0, \qquad P_{z^*}^{(I,2,2n)}(z, 0, t) = 0, \qquad n = 1, 2, \cdots$$

and such that for $\delta_0 \leqq |t| \leqq \delta_1$, $|s| \leqq s_0$, *and*

$$\frac{|z/r_1|}{|1-z/r_1|^3|1-z^*/r_1|} < \frac{\delta_0}{144\rho^2 s_0^2},$$

the series

$$E^{(I,k)}(z, z^*, t, s) = \frac{P^{(I,k,0)}(z, z^*)}{t} + \sum_{n=1}^{\infty} s^{2n} z^n P^{(I,k,2n)}(z, z^*, t), \quad k = 1, 2,$$

converges uniformly and absolutely.

Remark. Similarly for z and z^* sufficiently small there exist two linearly independent solutions of $\mathbf{T}^*(E)=0$, and these will be denoted by $E^{(II,1)}(z, z^*, t, s)$ and $E^{(II,2)}(z, z^*, t, s)$. The form of their series expansion is

$$E(z, z^*, t, s)=\frac{P^{(0)}(z, z^*)}{t}+\sum_{n=1}^{\infty} s^{2n}z^{*n}P^{(2n)}(z, z^*, t)$$

and the corresponding $P^{(2n)}$ satisfy the conditions

$$\begin{aligned} &P^{(II,1,0)}(0, z^*)=1, \qquad P_z^{(II,1,0)}(0, z^*)=0,\\ &P^{(II,1,2n)}(0, z^*, t)=0, \quad P_z^{(II,1,2n)}(0, z^*,t)=0, \end{aligned} \tag{2.21}$$

and

$$\begin{aligned} &P^{(II,2,0)}(0, z^*)=0, \qquad P_z^{(II,2,0)}(0, z^*)=1,\\ &P^{(II,2,2n)}(0, z^*, t)=0, \quad P_z^{(II,2,2n)}(0, z^*, t)=0. \end{aligned} \tag{2.22}$$

The existence of four generating functions will imply that the integral operator

$$\begin{aligned} \frac{1}{2\pi i}\sum_{k=1}^{2}\oint_{|\tau-\tau|=\delta}\int_{-1}^{1}\Big[&E^{(I,k)}(z, z^*, \tau-t, s)f_k\Big(\frac{z}{2}(1-s^2), \tau\Big)\\ &+E^{(II,k)}(z, z^*, \tau-t, s)g_k\Big(\frac{z^*}{2}(1-s^2), \tau\Big)\Big]\frac{ds\,d\tau}{(1-s^2)^{1/2}} \end{aligned} \tag{2.23}$$

is a solution of (2.1). Finally it should be noted that if the coefficients in (2.1) are entire functions of z and z^*, then $E^{(I,k)}$ and $E^{(II,k)}$, $k=1, 2$, are entire functions of z, z^*, s, and t except for an essential singularity at $t=0$.

For the existence of the integral operator, it is necessary that the generating functions satisfy the continuity conditions given by (2.5) and (2.6). Using (2.7) and (2.10a) we see that the first condition has the form

$$z^{-1}s^{-1}\mathbf{D}_1(E)=s\sum_{n=0}^{\infty} s^{2n}z^n\mathbf{D}_1(P^{(2n)}).$$

This is clearly a continuous function of s, z, z^*, and t for $|s|\leqq 1$, z and z^* in a neighborhood of the origin and $\delta_0<|t|<\delta_1$, $\delta_0>0$. Equation (2.6) takes the form

$$\begin{aligned} \frac{1}{zs}\Big[&\frac{-2}{t}(P^{(0)}_{zz^*z^*}+MP^{(0)}_z)+\frac{1}{zt}(P^{(0)}_{z^*z^*}+MP^{(0)})\\ &-\tfrac{1}{2}(P^{(2)}_{z^*z^*}+MP^{(2)})+\frac{1}{2tzs^2}(P^{(0)}_{z^*z^*}+MP^{(0)})-\frac{1}{t}AP^{(0)}-\frac{1}{t}LP^{(0)}_{z^*}\Big]+(\cdots), \end{aligned}$$

where $(\cdots)$ represents a continuous function. Equation (2.10a) implies the second and fourth terms equal zero. Equation (2.10b) implies the remaining terms in the bracket equal zero, leaving only a function continuous in the required region. We can now conclude that the integral operator defined by (2.23) exists and maps functions analytic in some neighborhood of the origin in $\mathbb{C}^2$ into the class of complex valued solutions of (2.1).

3. Invertibility of the integral operator and the time-independent case. In this section we will show that the integral operator (2.23) is an onto map for regular complex valued solutions. Next we will demonstrate that for solutions which are independent of t, equation (2.23) reduces to the solution given by Bergman [1].

Let $\tilde{U}(z, z^*, t)$ be a solution of (2.1) which is analytic for z and z^* sufficiently small and t in a neighborhood of the origin in $\mathbb{C}^1$. And let $U(z, z^*, t)$ be a solution of (2.1) derived from the integral operator (2.23). We shall establish that f_k and g_k can be chosen in (2.23) in such a way that U, U_z, and U_{z^*} assume preassigned analytic data on $z = 0$ and $z^* = 0$, namely the values of $\tilde{U}$ and its derivative there. Knowing the form of the E functions, using the initial conditions (2.19) and (2.20), and adding the assumption

$$g_k(0, t) = \frac{\partial g_k(0, t)}{\partial t} = 0, \qquad k = 1, 2, \tag{3.1}$$

we may represent $U(z, 0, t)$, $U_{z^*}(z, 0, t)$, $U(0, z^*, t)$, and $U_z(0, z^*, t)$ in terms of f_k and g_k, $k = 1, 2$. We obtain

$$U(z, 0, t) = \int_{-1}^{1} f_1\left(\frac{z}{2}(1-s^2), t\right)\frac{ds}{(1-s^2)^{1/2}}, \tag{3.2}$$

$$\begin{aligned} U(0, z^*, t) = &\int_{-1}^{1} g_1\left(\frac{z^*}{2}(1-s^2), t\right)\frac{ds}{(1-s^2)^{1/2}} + P^{(I,1,0)}(0, z^*)U(0, 0, t) \\ &+ P^{(I,2,0)}(0, z^*)U_{z^*}(0, 0, t). \end{aligned} \tag{3.3}$$

Similarly, expressions for $U_{z^*}(z, 0, t)$ and $U_z(0, z^*, t)$ can be obtained. The transformation

$$g(z, t) = \int_{-1}^{1} f\left(\frac{z}{2}[1-s^2]\right)\frac{ds}{(1-s^2)^{1/2}} \tag{3.4}$$

and its inverse

$$f(z/2, t) = -\frac{1}{2\pi}\int_{-1}^{1} g(z[1-s^2], t)\frac{ds}{s^2}, \tag{3.5}$$

(see [10, p. 114]) allow one to solve integral equations (3.2) and (3.3) for f_1 and g_1. Also the equations for $U_{z^*}(z, 0, t)$ and $U_z(0, z^*, t)$ could be solved for f_2 and g_2. With these ideas in mind, we are motivated to define four analytic functions from the given solution $\tilde{U}(z, z^*, t)$.

Define

$$f_1(z/2, t) \equiv -\frac{1}{2\pi}\int_{-1}^{1} \tilde{U}(z(1-s^2), 0, t)\frac{ds}{s^2}, \tag{3.6}$$

$$\begin{aligned} g_1(z^*/2, t) \equiv -\frac{1}{2\pi}\int_{-1}^{1} &[\tilde{U}(0, z^*(1-s^2), t) \\ &- P^{(I,1,0)}(0, z^*(1-s^2))\tilde{U}(0, 0, t) \\ &- P^{(I,2,0)}(0, z^*(1-s^2))\tilde{U}_{z^*}(0, 0, t)]\frac{ds}{s^2} \end{aligned} \tag{3.7}$$

with f_2 and g_2 defined in a similar manner. Clearly these functions are analytic and g_1 and g_2 satisfy (3.1).

Let $U(z, z^*, t)$ be the solution obtained from the integral operator (2.23) using the analytic functions f_k, g_k defined above. We claim that $U(z, z^*, t) = \tilde{U}(z, z^*, t)$. The definition of f_1, equation (3.6) and the transformation (3.4) imply

$$\tilde{U}(z, 0, t) = \int_{-1}^{1} f_1\left(\frac{z}{2}(1-s^2), t\right)\frac{ds}{(1-s^2)^{1/2}}.$$

Hence $U(z, 0, t) = \tilde{U}(z, 0, t)$. Similarly it may be shown that

$$\begin{aligned} U_{z^*}(z, 0, t) &= \tilde{U}_{z^*}(z, 0, t), \\ U(0, z^*, t) &= \tilde{U}(0, z^*, t), \\ U_z(0, z^*, t) &= \tilde{U}_z(0, z^*, t). \end{aligned} \tag{3.8}$$

Since both $\tilde{U}$ and U are analytic solutions in a neighborhood of the origin, they have a power series expansion say,

$$U(z, z^*, t) = \sum a_{mnp} z^m z^{*n} t^p, \qquad \tilde{U}(z, z^*, t) = \sum b_{mnp} z^m z^{*n} t^p$$

where sums are taken from $m, n, p = 0$ to ∞. Thus the conditions (3.8) imply

$$\begin{aligned} a_{m0p} &= b_{m0p}, & a_{0np} &= b_{0np} \\ a_{m1p} &= b_{m1p}, & a_{1np} &= b_{1np} \end{aligned} \qquad m, n, p = 0, 1, 2, \cdots. \tag{3.9}$$

Since U satisfies equation (2.1) we substitute the series expansion into this equation and set the coefficients of like powers of $z^m z^{*n} t^p$ equal to zero. This yields a relationship on the coefficients by which all the coefficients a_{mnp} can be uniquely expressed in terms of a_{0mp}, a_{1np}, a_{m0p}, and a_{m1p}. Replacing U by $\tilde{U}$ in the above argument, we find that the coefficients b_{mnp} satisfy the same relationship and can be uniquely expressed in terms of b_{0mp}, b_{1np}, b_{m0p}, b_{m1p}. Hence by (3.9), $U = \tilde{U}$. We summarize our results in the following theorem.

THEOREM 1. *For $|z|$ and $|z^*|$ sufficiently small, $|s| \leqq 1$ and $\delta_0 < |t| < \delta_1$, where $\delta_0 > 0$ and δ_1 is arbitrarily large, there exist four functions $E^{(j,k)}(z, z^*, t, s)$, $j = I, II$, $k = 1, 2$, such that in some neighborhood of the origin*

$$U(z, z^*, t) = \frac{1}{2\pi i} \sum_{k=1}^{2} \oint_{|\tau - t| = \delta} \int_{-1}^{1} \left[E^{(I,k)}(z, z^*, \tau - t, s) f_k\left(\frac{z}{2}(1-s^2), \tau\right) + E^{(II,k)}(z, z^*, \tau - t, s) g_k\left(\frac{z^*}{2}(1-s^2), \tau\right)\right] \cdot \frac{ds\, d\tau}{(1-s^2)^{1/2}}$$

is an analytic solution of (2.1) *where f_k and g_k are arbitrary analytic functions defined in a neighborhood of the origin in $\mathbb{C}^2$. Conversely, if $U(z, z^*, t)$ is a complex valued analytic solution of* (2.1) *defined in some neighborhood of the origin, then U can be represented by* (2.23) *where the functions f_k and g_k are given by* (3.6)–(3.7) *and are analytic in some neighborhood of the origin in $\mathbb{C}^2$.*

Since the operator defined by (2.23) is based upon the work of Bergman, it is not surprising that it reduces to Bergman's operator [1, p. 620] when $U(z, z^*, t) = U(z, z^*)$ is independent of t. Equation (2.1) becomes

$$U_{zzz^*z^*} + MU_{zz} + LU_{zz^*} + NU_{z^*z^*} + AU_z + BU_{z^*} + CU = 0$$

and (3.6) and (3.7) imply that the associate functions f_1 and g_1 are independent of t. This is also true for f_2 and g_2. Thus it is possible to integrate termwise in (2.3) with respect to τ. We define $\tilde{P}^{(2n)}(z, z^*)$ as

$$\tilde{P}^{(2n)}(z, z^*) \equiv \frac{1}{2\pi i} \oint_{|\tau - t| = \delta} P^{(2n)}(z, z^*, \tau - t)\, d\tau.$$

We find that $\tilde{P}^{(0)}$ satisfies (2.10a), and $\tilde{P}^{(2n)}$ are defined recursively by

$$\tilde{P}^{(2)}_{z^*z^*} + M\tilde{P}^{(2)} = -2[A\tilde{P}^{(0)} + L\tilde{P}^{(0)}_{z^*} + 2\tilde{P}^{(0)}_{zz^*z^*} + 2M\tilde{P}^{(0)}_{z}] \tag{3.10}$$

and

$$\begin{aligned} \tilde{P}^{(2n+4)}_{z^*z^*} + M\tilde{P}^{(2n+4)} = {} & \frac{-1}{n^2 + 2n + 3/4} \\ & \cdot [(2n+1)\tilde{P}^{(2n+2)}_{zz^*z^*} + (n + 1/2)(L\tilde{P}^{(2n+2)}_{z^*} + A\tilde{P}^{(2n+2)}) \\ & + (2n+1)M\tilde{P}_z^{(2n+2)} + N\tilde{P}^{(2n)}_{z^*z^*} + B\tilde{P}^{(2n)}_{z^*} + C\tilde{P}^{(2n)} \\ & + A\tilde{P}^{(2n)}_{z} + \tilde{P}^{(2n)}_{zzz^*z^*} + M\tilde{P}^{(2n)}_{zz} + L\tilde{P}^{(2n)}_{zz^*}], \qquad n = 0, 1, 2, \cdots. \end{aligned} \tag{3.11}$$

The representation for U becomes

$$U(z, z^*) = \int_{-1}^{1} E(z, z^*, s) f\left(\frac{z}{2}(1 - s^2)\right) \frac{ds}{(1 - s^2)^{1/2}}$$

where

$$E(z, z^*, s) = \tilde{P}^{(0)}(z, z^*) + \sum_{n=1}^{\infty} s^{2n} z^n \tilde{P}^{(2n)}(z, z^*)$$

with the $p^{(2n)}$, $n = 0, 1, \cdots$, satisfying equations (2.10a), (3.10) and (3.11) respectively. A comparison of these equations with Bergman's operator for fourth order elliptic equations shows that they are identical.

4. Fourth order radial equations in two space variables. We will now use the theorem of the previous section to construct an integral operator that maps harmonic functions with a complex parameter onto real solutions of the differential equation

$$\Delta^2_{p+2} u(\mathbf{x}, t) + A(r^2)\Delta_{p+2} u(\mathbf{x}, t) + B(r^2) u(\mathbf{x}, t) = C(r^2) u_t(x, t) \tag{4.1}$$

where $\mathbf{x} = (x_1, \cdots, x_{p+2})$, $r^2 = x_1^2 + x_2^2 + \cdots + x_{p+2}^2$, u_t is the partial derivative with respect to t, and $A(r^2)$, $B(r^2)$, $C(r^2)$ are real valued analytic functions of r^2 for $|r| \leqq R$. We first consider the case when $p = 0$. In terms of the complex variables z

and z^*, the equation takes the form

$$16U_{zzz^*z^*}+4A(zz^*)U_{zz^*}+B(zz^*)U-C(zz^*)U_t=0 \tag{4.2}$$

where $U(z, z^*, t)=u((z+z^*)/2, (z-z^*)/(2i), t)=u(x, y, t)$ and the coefficients are analytic in some neighborhood of the origin in $\mathbb{C}^2$. This is a special case of (2.1) in which the coefficients are real valued when $z^*=\bar{z}$ and the coefficients M, N, A, B equal zero. Complex valued solutions are provided by (2.23), but the generating functions $E^{(I,k)}(z, z^*, t, s)$, $k=1, 2$, now satisfy the equation

$$\begin{aligned} &z^{-1}s^{-1}(1-s^2)[E_{zz^*z^*s}+\tfrac{1}{8}A(zz^*)E_{sz^*}]+\tfrac{1}{4}z^{-2}s^{-2}(1-2^2)^2E_{z^*z^*ss}\\ &-z^{-1}s^{-1}[E_{zz^*z^*}+\tfrac{1}{8}A(zz^*)E_{z^*}]-\tfrac{3}{4}z^{-2}s^{-3}(1-s^4)E_{sz^*z^*}\\ &+\tfrac{3}{4}z^{-2}s^{-4}E_{z^*z^*}+E_{zzz^*z^*}+\tfrac{1}{4}A(zz^*)E_{zz^*}+\tfrac{1}{16}B(zz^*)E\\ &-\tfrac{1}{16}C(zz^*)E_t=0. \end{aligned} \tag{4.3}$$

The form of $E^{(I,k)}(z, z^*, t, s)$ is given in (2.7), but the recursion equations for the $p^{(2n)}(z, z^*, t)$ are simpler. These equations become

$$P^{(0)}_{z^*z^*}=0, \tag{4.4a}$$

$$P^{(2)}_{z^*z^*}=-\frac{2}{t}\left[\frac{1}{4}A(zz^*)P^{(0)}_{z^*}+2P^{(0)}_{zz^*z^*}\right], \tag{4.4b}$$

$$\begin{aligned} P^{(2n+4)}_{z^*z^*}=-\frac{1}{n^2+2n+3/4}\Big[&(2n+1)P^{(2n+2)}_{zz^*z^*}+\frac{1}{4}\left(n+\frac{1}{2}\right)A(zz^*)P^{(2n+2)}_{z^*}\\ &+\frac{1}{16}B(zz^*)P^{(2n)}+P^{(2n)}_{zzz^*z^*}+\frac{1}{4}A(zz^*)P^{(2n)}_{zz^*}-\frac{1}{16}C(zz^*)P^{(2n)}_t\Big] \end{aligned} \tag{4.4c}$$

$$n=0,\ 1,\ 2,\cdots.$$

The initial conditions (2.19)–(2.21) imply that $p^{(1,0)}(z, z^*)/t=1/t$ and $p^{(2,0)}(z, z^*)/t=z^*/t$.

It should be noted that when $z^*=\bar{z}$ and s and t are real valued then

$$E^{(II,k)}(z, \bar{z}, t, s)=\overline{E^{(I,k)}(z, \bar{z}, t, s)}\equiv\bar{E}^{(I,k)}(\bar{z}, z, t, s), \qquad k=1, 2, \tag{4.5}$$

since the complex conjugate of (4.3) becomes precisely the equation that $E^{(II,k)}(z, z^*, t, s)$ satisfies. Not only the form of the power series for $\bar{E}^{(I,k)}$ is the same as that for $E^{(II,k)}$ but their initial conditions also agree. Thus $P^{(II,k,2n)}(z, \bar{z}, t)$ and $\bar{p}^{(I,k,2n)}(\bar{z}, z, t)$ must satisfy the same recursion equations and (4.5) follows. Furthermore, it is easily verified that the generating function $E^{(I,1)}$ is a real function of $r^2=z\bar{z}$, t, and s, and we define the function $E^{(1)}(r^2, t, s)$ by

$$E^{(1)}(r^2, t, s)\equiv E^{(I,1)}(z, z^*, t, s). \tag{4.6}$$

Also $E^{(I,2)}(z, \bar{z}, t, s)$ is of the form

$$E^{(I,2)}(z, z^*, t, s)=z^*\tilde{E}^{(2)}(r^2, t, s) \tag{4.7}$$

where $\tilde{E}^{(2)}(r^2, t, s)$ is a real function of r^2, t and s (cf. [3] or [6, p. 64]). The function

$E^{(1)}(r^2, t, s)$ has a series expansion of the form

$$E^{(1)}(r^2, t, s) = \frac{1}{t} + \sum_{n=0}^{\infty} s^{2n} e_n^{(1)}(r^2, t) \tag{4.8}$$

with

$$e_n^{(1)}(0, t) = \frac{\partial e_n(0, t)}{\partial r^2} = 0, \qquad n = 1, 2, \cdots. \tag{4.9}$$

From the results of § 2, it follows that this series converges absolutely and uniformly for r in some neighborhood of the origin, $\delta_0 < t < \delta_1$ where $\delta_0 > 0$ and δ_1 is arbitrarily large, and $|s| \leqq 1$. The second generating function converges in the same region and has the expansion

$$\tilde{E}^{(2)}(r^2, t, s) = \frac{1}{t} + \sum_{n=1}^{\infty} s^{2n} \tilde{e}_n^{(2)}(r^2, t) \tag{4.10}$$

with

$$\tilde{e}_n^{(2)}(0, t) = \frac{\partial \tilde{e}_n^{(2)}(0, t)}{\partial r^2} = 0, \qquad n = 1, 2, \cdots. \tag{4.11}$$

THEOREM 2. *Let $H^{(1)}(x_1, x_2, \tau)$, $H^{(2)}(x_1, x_2, \tau)$ be arbitrary harmonic functions with a complex parameter τ, defined for (x_1, x_2) in a starlike domain with respect to the origin and τ in a neighborhood of the origin in $\mathbb{C}^1$. Then in some neighborhood of the origin, the function defined by*

$$\begin{aligned} u(x_1, x_2, t) = \frac{1}{2\pi i} \oint_{|\tau - t| = \delta} \int_{-1}^{1} & [E^{(1)}(r^2, \tau - t, s) H^{(1)}(\mathbf{x}(1-s^2), \tau) \\ & + E^{(2)}(r^2, \tau - t, s) \\ & \cdot H^{(2)}(\mathbf{x}(1-s^2), \tau)] \frac{ds\, d\tau}{(1-s^2)^{1/2}} \end{aligned} \tag{4.12}$$

is a real valued solution of (4.1) *for $p = 0$. The function $E^{(1)}(r^2, t, s)$ has the expansion given by* (4.8) *while $E^{(2)}(r^2, t, s)$ has the uniformly convergent expansion*

$$\begin{aligned} E^{(2)}(r^2, t, s) &= r^2[1 - s^2]\left(\frac{1}{t} + \sum_{n=1}^{\infty} s^{2n} \tilde{e}_n^{(2)}(r^2, t)\right) \\ &\equiv \frac{r^2}{t} + \sum_{n=1}^{\infty} s^{2n} e_n^{(2)}(r^2, t). \end{aligned} \tag{4.13}$$

Furthermore, every real valued analytic solution of (4.1) *(for $p = 0$) in a neighborhood of the origin can be represented by* (4.12).

Proof. In (2.23), let $g_1(z^*/2, t) = \bar{f}_1(z^*/2, t)$, $g_2(z^*/2, t) = \bar{f}_2(z^*/2, t)$, and $z^* = \bar{z}$. Then by using (4.5), it can be shown that the integral operator (2.23) equals the expression

$$
(4.14)\quad 2\operatorname{Re}\left(\frac{1}{2\pi i}\oint_{|\tau-t|=\delta}\int_{-1}^{1}\left[E^{(I,1)}(z,z^*,\tau-t,s)f_1\left(\frac{z}{2}(1-s^2),\tau\right)\right.\right.
$$
$$
\left.\left.+E^{(I,2)}(z,z^*,\tau-t,s)f_2\left(\frac{z}{2}(1-s^2),\tau\right)\right]\frac{ds\,d\tau}{(1-s^2)^{1/2}}\right),
$$

where "Re" denotes "take the real part." Thus, pairs of analytic functions are mapped onto real valued solutions. On the other hand, using the facts (4.5)–(4.7), equation (2.23) equals (4.12) if we set

$$
H^{(1)}(x_1,x_2,t)=\frac{1}{2}\left[f_1\left(\frac{z}{2},t\right)+\bar{f}_1\left(\frac{\bar{z}}{2},t\right)\right],
$$

$$
r^2H^{(2)}(x_1,x_2,t)=\frac{1}{2}\left[\bar{z}f_2\left(\frac{z}{2},t\right)+z\bar{f}_2\left(\frac{\bar{z}}{2},t\right)\right].
$$

(Recall from (3.1) that we may assume $f_2(0,t)=0$). Finally if $f_1(z,t)$ has the series expansion

$$
f_1\left(\frac{z}{2},t\right)=\sum_{m,n}^{\infty}a_{mn}\left(\frac{z}{2}\right)^m t^n,
$$

then

$$
H^{(1)}(x_1,x_2,t)=\sum_{m,n=0}^{\infty}\left[\operatorname{Re}a_{mn}\left(\frac{z}{2}\right)^m\right]t^n.
$$

Therefore $H^{(1)}(x_1,x_2,t)$ is a harmonic function in (x_1,x_2) with a complex parameter t and the theorem follows.

The differential equation (4.3) which $E(z,z^*,t,s)$ satisfies is now transformed for both functions $E^{(1)}(r^2,t,s)$ and $E^{(2)}(r^2,t,s)$ into the differential equation

$$
\begin{aligned}
&\frac{1-s^2}{rs}\left[2E_{rrrs}-\frac{2}{r}E_{rrs}-\frac{2}{r^2}E_{rs}+A(r^2)E_{rs}\right]\\
(4.15)\quad &+\frac{(1-s^2)^2}{r^2s^2}\left[E_{rrss}-\frac{1}{r}E_{rss}\right]-\frac{1}{rs^2}\left[2E_{rrr}+\frac{2}{r}E_{rr}-\frac{2}{r^2}E_r+A(r^2)E_r\right]\\
&-\frac{3(1-s^4)}{s^3r^2}\left[E_{rrs}-\frac{1}{r}E_{rs}\right]+\frac{3}{s^4r^2}\left[E_{rr}-\frac{1}{r}E_r\right]+E_{rrrr}+\frac{2}{r}E_{rrr}\\
&-\frac{1}{r^2}E_{rr}+\frac{1}{r_3}E_r+A(r^2)\left[E_{rr}+\frac{1}{r}E_r\right]+B(r^2)E-C(r^2)E_t=0
\end{aligned}
$$

and the initial conditions

$$
(4.16)\quad \begin{aligned}
&E^{(1)}(0,t,s)=1/t, && E^{(1)}_{r^2}(0,t,s)=0,\\
&E^{(2)}(0,t,s)=0, && E^{(2)}_{r^2}(0,t,s)=(1-s^2)/t.
\end{aligned}
$$

These equations follow from (4.9) and (4.11).

There is one more representation for solutions in two variables which will prove valuable in our attempt to represent solutions with n variables. This type of representation was originally used by Gilbert in [9] for the elliptic equation $\Delta_2 u + c(r^2)u = 0$ and later extended by Colton and Gilbert in [6] to $\Delta_2^2 u + A(r^2)\Delta u + B(r^2)u = 0$.

THEOREM 3. *Let $h^{(1)}(x_1, x_2, \tau)$ and $h^{(2)}(x_1, x_2, \tau)$ be arbitrary harmonic functions, with a complex parameter τ, defined for (x_1, x_2) in some disk centered at the origin and τ in a neighborhood of the origin. Then the function*

$$\begin{aligned} u(x_1, x_2, t) = {} & h^{(1)}(x_1, x_2, t) + r^2 h^{(2)}(x_1, x_2, t) \\ & + \frac{1}{2\pi i}\oint_{|\tau - t| = \delta} \int_0^1 [\sigma G^{(1)}(r^2, \tau - t, 1-\sigma^2) h^{(1)}(x_1\sigma^2, x_2\sigma^2, \tau) \\ & + \sigma G^{(2)}(r^2, \tau - t, 1-\sigma^2) \\ & \cdot h^{(2)}(x_1\sigma^2, x_2\sigma^2, \tau)]\, d\sigma\, d\tau \end{aligned} \tag{4.17}$$

where

$$G^{(1)}(r^2, t, u) = \sum_{n=1}^{\infty} \frac{2e_n^{(i)}(r^2, t)\Gamma(n+\frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(n)} u^{n-1}, \qquad i = 1, 2, \tag{4.18}$$

is a solution to (4.1) *for $p = 0$. Furthermore, every real valued analytic solution of* (4.1) *for $p = 0$ has a representation of the form* (4.17).

Proof. The harmonic functions defined in Theorem 2, $H^{(1)}(x_1, x_2, \tau)$ and $H^{(2)}(x_1, x_2, \tau)$ can be expanded in some disk about the origin in the following form:

$$H(x_1, x_2, \tau) = \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} b_{mn} r^{|m|} e^{im\theta} \tau^n \tag{4.19}$$

where $\bar{b}_{mn} = b_{-mn}$. Define the harmonic functions $h^{(1)}(x_1, x_2, \tau)$ and $h^{(2)}(x_1, x_2, \tau)$ by

$$h^{(k)}(x_1, x_2, \tau) = \int_{-1}^{1} H^{(k)}(x_1(1-s^2), x_2(1-s^2), \tau)\frac{ds}{(1-s^2)^{1/2}}, \quad k = 1, 2. \tag{4.20}$$

A standard formula for the beta function together with (4.19) implies that in a disk about the origin, $h^{(1)}(x_1, x_2, \tau)$ and $h^{(2)}(x_1, x_2, \tau)$ have expansions of the form

$$h(x_1, x_2, \tau) = \sum_{n=0}^{\infty} \sum_{m=-\infty}^{\infty} b_{mn} r^{|m|} e^{im\theta} \frac{\Gamma(\frac{1}{2})\Gamma(|m|+\frac{1}{2})}{\Gamma(|m|+1)} \tau^m. \tag{4.21}$$

We substitute the series expansions for $E^{(1)}$ and $E^{(2)}$ into (4.12) and integrate termwise with respect to s. Also we perform termwise integration of the kernel function expansion in (4.17). The theorem now follows by making use of (4.21) and comparing the terms of the series expansions.

Remark. We note that an important consequence of an invertible integral operator which acts on analytic or harmonic functions is that it can be used to obtain complete families of solutions. This, in turn, has been a useful tool for approximating solutions to standard boundary value problems (see [8]). In the present context, a complete family of solutions to (4.1) in a compact starlike domain can be obtained from the integral operator (4.17) by using Runge's theorem [7, p. 47] and the completeness of the harmonic polynomials [14]. In particular, we let $\{h_n(x_1, x_2)\}_{n=0}^{\infty}$ denote the set of harmonic polynomials and in (4.17) set $h(x_1, x_2, t) = h_m(x_1, x_2)t^n$ where $m, n = 0, 1, 2, \cdots$. Also it can be shown that the integral operator (4.14) yields a complete family of real valued solutions to (1.1) in a starlike domain by setting $f(z, t) = z^m t^n$, $m, n = 0, 1, 2, \cdots$.

5. The method of ascent to $p+2$ space variables. We now consider (4.1) for $p>0$. The remarkable fact about the method of ascent to solutions for higher dimensions is that the kernels in the n-dimensional case are the same as $G^{(1)}(r^2, t, u)$ and $G^{(2)}(r^2, t, u)$ in the two dimensional case. The integral operator is merely modified by a factor of σ^p and the harmonic functions are functions of $p+3$ variables, $x_1, x_2, \cdot\cdot, x_{p+2}, t$. This result is analogous to the results for elliptic equations (cf. [6], [9]). The coefficients of (4.1) are assumed to be analytic functions for $r<R$.

THEOREM 4. *Let $E^{(1)}(r^2, t, s; p)$ and $E^{(2)}(r^2, t, s; p)$ be solutions of the differential equation*

$$\begin{aligned}(5.1)\quad &\frac{1-s^2}{rs}\left[2E_{rrrs} + \frac{(2-2p)}{r}E_{rrs} + \frac{2(p-1)}{r^2}E_{rs} + A(r^2)E_{rs}\right] \\ &+\frac{(1-s^2)^2}{r^2s^2}\left[E_{rrss} - \frac{1}{r}E_{rss}\right] + \frac{p-1}{rs^2}\left[2E_{rrr} + \frac{2}{r}E_{rr} - \frac{2}{r^2}E_r + A(r^2)E_r\right] \\ &+\frac{(2p-3)(1-s^4)}{r^2s^3}\left[E_{rrs} - \frac{1}{r}E_{rs}\right] + \frac{(p-3)(p-1)}{r^2s^4}\left[E_{rr} - \frac{1}{r}E_r\right] + E_{rrrr} \\ &+\frac{2}{r}E_{rrr} - \frac{1}{r^2}E_{rr} + \frac{1}{r^3}E_r + A(r^2)\left[E_{rr} + \frac{1}{r}E_r\right] + B(r^2)E - C(r^2)E_t = 0\end{aligned}$$

which are regular for r in some neighborhood of the origin, $|s| \leqq 1$, and $\delta_0 < t < \delta_1$ where δ_0 is positive and δ_1 is arbitrarily large. Suppose they satisfy the boundary conditions

$$(5.2)\qquad E^{(1)}(0, t, s; p) = \frac{1}{t}, \qquad E^{(1)}_{r^2}(0, t, s; p) = 0,$$

$$(5.3)\qquad E^{(2)}(0, t, s; p) = 0, \qquad E^{(2)}_{r^2}(0, t, s; p) = \frac{1}{t}\left(1 - \frac{s^2}{p+1}\right).$$

Let $H^{(1)}(\mathbf{x}, \tau)$ and $H^{(2)}(\mathbf{x}, \tau)$ be harmonic functions of $\mathbf{x} = (x_1, \cdots, x_n)$ with a complex parameter τ defined for $\mathbf{x}$ in a starlike region with respect to the origin and τ in a neighborhood of the origin in $\mathbb{C}^1$. Then for $\delta_0 < \delta < \delta_1$,

$$\begin{aligned}(5.4)\quad u(x, t) = \frac{1}{2\pi i}\oint_{|\tau - t| = \delta}\int_{-1}^{1} &s^p[E^{(1)}(r^2, \tau - t, s; p)H^{(1)}(\mathbf{x}(1-s^2), \tau) \\ &+ E^{(2)}(r^2, \tau - t, s; p) \\ &\cdot H^{(2)}(\mathbf{x}(1-s^2), \tau)]\frac{ds\, d\tau}{(1-s^2)^{1/2}}\end{aligned}$$

is a solution of (4.1) *in a neighborhood of the origin.*

Proof. The differential equation for the E function can be verified directly by substituting the integral operator (5.4) into (4.1) and integrating by parts. The choice of these boundary conditions is for the purpose of normalization in the method of ascent. We observe that for $p=0$, the differential equation and boundary conditions reduce to (4.15) and (4.16).

THEOREM 5. *There exist functions* $E^{(1)}(r^2, t, s; p)$ *and* $E^{(2)}(r^2, t, s; p)$ *which satisfy the differential equation* (5.1) *and boundary conditions* (5.2) *and* (5.3). *They have the expansions*

$$E^{(1)}(r^2, t, s; p) = \frac{1}{t} + \sum_{n=1}^{\infty} s^{2n} e_n^{(1)}(r^2, t; p), \tag{5.5}$$

$$E^{(2)}(r^2, t, s; p) = \frac{r^2}{t} + \sum_{n=1}^{\infty} s^{2n} e_n^{(2)}(r^2, t; p), \tag{5.6}$$

which converge uniformly and absolutely for r *in a neighborhood of the origin,* $\delta_0 < t < \delta_1$ *where* $\delta_0 > 0$ *and* δ_1 *is arbitrarily large, and* $|s| \leqq 1$.

Proof. Setting $e_0^{(1)}(r^2, t; p) = 1/t$ and $e_0^{(2)}(r^2, t; p) = r^2/t$ and substituting the series (5.5) and (5.6) into (5.1) we find that both $e_n^{(1)}(r^2, t; p)$ and $e_n^{(2)}(r^2, t; p)$ for $n \geqq 0$ satisfy the recursion equation:

$$\begin{aligned}
&(2n+p+3)(2n+p+1)\left(e''_{n+2} - \frac{e'_{n+2}}{r}\right) + 2(2n+p+1)re'''_{n+1} \\
&-(4n+2)(2n+p+1)\left(e''_{n+1} - \frac{e'_{n+1}}{r}\right) + (2n+p+1)rA(r^2)e'_{n+1} \\
&+r^2 e''''_n + r(2-4n)e'''_n + (4n^2-1)\left(e''_n - \frac{e'_n}{r}\right) \\
&+r^2 A(r^2)\left(e''_n + \frac{1-2n}{r} e'_n\right) + r^2 B(r^2) e_n - r^2 C(r^2)\frac{\partial e_n}{\partial t} = 0,
\end{aligned} \tag{5.7}$$

with

$$e_n(0, t; p) = \frac{\partial e_n(0, t; p)}{\partial r^2} = 0, \qquad n = 2, 3, \cdots. \tag{5.8}$$

Here $'$ denotes differentiation with respect to r. The functions $e_1^{(1)}$ and $e_1^{(2)}$ must satisfy the equations

$$e_1^{(1)''} - \frac{e_1^{(1)'}}{r} = 0 \tag{5.9}$$

and

$$e_1^{(2)''} - \frac{e_1^{(2)'}}{r} = \frac{-2r^2}{(p+1)t} A(r^2) \tag{5.10}$$

respectively. For $e_1^{(1)}$, we choose the boundary conditions

$$e_1^{(1)}(0, t; p) = \frac{\partial e_1^{(1)}(0, t; p)}{\partial r^2} = 0 \tag{5.11}$$

which implies that

$$e_1^{(1)}(r^2, t; p) = 0. \tag{5.12}$$

In solving for $e_1^{(2)}$, it will be desirable to choose constants in such a way that the term $1/(p+1)$ may be factored and the boundary condition (5.3) reduces to equation (4.16) when $p = 0$. The appropriate conditions are

$$e_1^{(2)}(0, t; p) = 0, \qquad \frac{\partial e_1^{(2)}(0, t; p)}{\partial r^2} = -\frac{1}{(p+1)t} \tag{5.13}$$

and thus

$$e_1^{(2)}(r^2, t; p) = \frac{-1}{(p+1)t}\left[r^2 + r^2\int_0^r \xi A(\xi^2)\, d\xi - \int_0^r \xi^3 A(\xi^2)\, d\xi\right]. \tag{5.14}$$

For $n \geqq 2$, $e_n^{(i)}(r^2, t; p)$, $i = 1, 2$, are defined recursively by (5.7) and (5.8) and then the series (5.5) and (5.6) formally satisfy (5.1) and the boundary conditions (5.2) and (5.3) respectively. In order to show that the series (5.5) and (5.6) converge, observe that $E^{(i)}(r^2, t, s)$ and $E^{(i)}(r^2, t, s; 0)$, $i = 1, 2$, satisfy the same differential equation and boundary conditions. Since their expansions have the same form, it follows that $e_n^{(i)}(r^2, t; 0) = e^{(i)}(r^2, t)$, $i = 1, 2$. Thus, the series (5.5) and (5.6) converge for r in a neighborhood of the origin, $\delta_0 < t < \delta_1$, $\delta_0 > 0$, and $|s| \leqq 1$ when $p = 0$. Next we define new functions $c_n^{(i)}(r^2, t; p)$, $i = 1, 2$, $n = 0, 1, 2, \cdots$, by the formulas

$$\begin{gathered} c_0^{(1)}(r^2, t; p) = 1/t, \qquad c_0^{(2)}(r^2, t; p) = r^2/t, \\ c_n^{(i)}(r^2, t; p) = \frac{2e_n^{(i)}(r^2, t; p)\Gamma(n + p/2 + 1/2)}{\Gamma(n)\Gamma(p/2 + 1/2)}, \qquad n = 1, 2, \cdots. \end{gathered} \tag{5.15}$$

Using (5.12) and (5.14), and the values of $e_0^{(1)}$ and $e_0^{(2)}$ in (5.7) with $n = 0$, we determine that $c_1^{(i)}$ and $c_2^{(i)}$ are given by the following formulas:

$$c_1^{(1)}(r^2, t; p) = 0, \tag{5.16}$$

$$\begin{aligned} c_1^{(2)}(r^2, t; p) &= (p+1)e_1^{(2)}(r^2, t; p) \\ &= -\frac{1}{t}\left[r^2 + r^2\int_0^r \xi A(\xi^2)\, d\xi - \int_0^r \xi^3 A(\xi^2)\, d\xi\right], \end{aligned} \tag{5.17}$$

$$\begin{gathered} c_2^{(1)\prime\prime} - \frac{1}{r}c_2^{(1)\prime} = \frac{r^2}{2}\left(\frac{1}{t^2}C(r^2) - \frac{1}{t}B(r^2)\right), \\ \left(c_2^{(2)\prime\prime} - \frac{1}{r}c_2^{(2)\prime}\right) + rc_1^{(2)\prime\prime} - \left(c_1^{(2)\prime\prime} - \frac{1}{r}c_1^{(2)\prime}\right) + \frac{r}{2}A(r^2)c_1^{(2)\prime} \\ = -\frac{r^2}{2}\left[2A(r^2) + \frac{r^2}{2}B(r^2) + \frac{r^2}{2t}C(r^2)\right]. \end{gathered} \tag{5.18}$$

For $n = 1, 2, \cdots$, $c_n^{(1)}(r^2, t; p)$, $c_n^{(2)}(r^2, t; p)$ both satisfy

$$\begin{aligned} &4n(n+1)\Big(c''_{n+2}-\frac{1}{r}c'_{n+2}\Big)+4rnc'''_{n+1}-4n(2n+1)\Big(c''_{n+1}-\frac{1}{r}c'_{n+1}\Big) \\ &\quad +2nrA(r^2)c'_{n+1}+r^2c''''_n+(2-4n)rc'''_n+(4n^2-1)\Big(c''_n-\frac{c'_n}{r}\Big) \\ &\quad +r^2A(r^2)\Big(c''_n+\frac{1-2n}{r}c'_n\Big)+r^2B(r^2)c_n-r^2C(r^2)\frac{\partial c_n}{\partial t}=0 \end{aligned} \tag{5.19}$$

with

$$c_n(0, t; p) = \frac{\partial c_n(0, t; p)}{\partial r^2} = 0, \qquad n = 2, 3, \cdots. \tag{5.20}$$

Since (5.16)–(5.20) do not involve p, they imply that the $c_n^{(i)}(r^2, t; p)$ are in fact independent of p. This result, together with the fact that the series (5.5) and (5.6) converge for $p = 0$ implies from (5.15) that these series converge for any positive integer p in the domain stated above. This concludes the theorem.

THEOREM 6. *Let $h^{(1)}(\mathbf{x}, \tau)$ and $h^{(2)}(\mathbf{x}, \tau)$ be arbitrary harmonic functions of $\mathbf{x} = (x_1, \cdots, x_{p+2})$ with complex parameter τ, defined for $\mathbf{x}$ in some sphere centered at the origin and τ in a neighborhood of the origin in $\mathbb{C}^1$. Then for $\delta_0 < \delta < \delta_1$,*

$$\begin{aligned} u(\mathbf{x}, t) &= h^{(1)}(\mathbf{x}, t) + r^2 h^{(2)}(\mathbf{x}, t), \\ &\quad + \frac{1}{2\pi i}\oint_{|\tau - t| = \delta}\int_0^1 \sigma^{p+1}[G^{(1)}(r^2, \tau - t, 1-\sigma^2)h^{(1)}(\mathbf{x}\sigma^2, \tau) \\ &\qquad + G^{(2)}(r^2, \tau - t, 1-\sigma^2)h^{(2)}(\mathbf{x}\sigma^2, \tau)]\, d\sigma\, d\tau \end{aligned} \tag{5.21}$$

is a solution of (4.1). *The functions $G^{(1)}$ and $G^{(2)}$ are independent of p, have expansions of the form*

$$G^{(k)}(r^2, t, u) = \sum_{n=1}^{\infty} c_n^{(k)}(r^2, t)u^{n-1}, \qquad k = 1, 2$$

with $c_n^{(k)}$ defined in (5.15), *and satisfy the boundary conditions*

$$G^{(1)}(0, t, u) = 0, \qquad G^{(1)}_{r^2}(0, t, u) = 0, \tag{5.22}$$

$$G^{(2)}(0, t, u) = 0, \qquad G^{(2)}_{r^2}(0, t, u) = -1/t. \tag{5.23}$$

Proof. After substituting (5.5) and (5.6) into (5.4) and defining

$$h^{(k)}(\mathbf{x}, \tau) = \int_0^1 s^p H^{(k)}\Big(\mathbf{x}(1-s^2), \tau\Big)\frac{ds}{(1-s^2)^{1/2}}, \qquad k = 1, 2,$$

the calculations follow exactly as in Theorem 3. A comparison of the expansions for the G functions above with those in Theorem 3 equation (4.18), shows that they are identical since $c_n(r^2, t; p)$ are independent of p. The boundary conditions follow from (5.20), (5.16) and (5.17).

In conclusion we wish to point out that it remains to be shown that the integral representation of Theorem 6 is invertible and yields a complete family of solutions for $p>0$. Also it should be noted that if the coefficients of (4.1) are entire functions then the kernel functions will be entire functions of r^2.

Acknowledgments. The author wishes to thank R. P. Gilbert and D. Colton for many valuable suggestions.

REFERENCES

[1] S. BERGMAN, *Solutions of linear partial differential equations of the fourth order*, Duke Math. J., 11 (1944), pp. 617–649.

[2] ———, *Integral operators in the theory of linear partial differential equations*, Springer, Berlin 1961.

[3] P. M. BROWN, Ph.D Dissertation, Indiana Univ., Bloomington, 1973.

[4] D. COLTON, *Bergman operators for parabolic equations in two space variables*, Proc. Amer. Math. Soc., 38 (1973), pp. 119–126.

[5] ———, *Cauchy's problem for a class of fourth order elliptic equations in two independent variables*, Applicable Anal., 1 (1971), pp. 13–22.

[6] D. COLTON AND R. P. GILBERT, *Integral operators and complete families of solutions for* $\Delta_{p+2}^2 u(\mathbf{x})+A(r^2)\Delta_{p+2}u(\mathbf{x})+B(r^2)u(\mathbf{x})=0$, Arch. Rational Mech. Anal., 43 (1971), pp. 62–78.

[7] B. A. FUKS, *Special Chapters in the Theory of Analytic Functions of Several Complex Variables*, American Mathematical Society, Providence, RI, 1965.

[8] R. P. GILBERT, *The construction of solutions for boundary value problems by function theoretic methods*, this Journal, 1 (1970), pp. 96–114.

[9] ———, *A method of ascent for solving boundary value problems*, Bull. Amer. Math. Soc., 75 (1969), pp. 1286–1289.

[10] ———, *Function Theoretic Methods in Partial Differential Equations*, Academic Press, New York, 1969.

[11] C. D. HILL, *A method for the construction of reflection laws for a parabolic equations*, Trans. Amer. Math. Soc., 133 (1968). pp. 357–372.

[12] R. LATTES AND J. L. LIONS, *The Method of Quasireversibility. Applications to Partial Differential Equations*, American Elsevier, New York, 1969.

[13] L. E. PAYNE, *Symposium on Non-Well-Posed Problems and Logarithmic Convexity*, Lecture Notes in Mathematics, Vol. 161, Springer-Verlag, Berlin, 1973.

[14] N. DU PLESSIS, *Runge's theorem for harmonic functions*, J. London Math. Soc., 1 (1969), pp. 404–408.

[15] I. N. VEKUA, *New Methods for Solving Elliptic Equations*, John Wiley, New York, 1967.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

ON AN ELLIPTIC BOUNDARY VALUE PROBLEM WITH MIXED BOUNDARY CONDITIONS, ARISING IN SUSPENDED SEDIMENT TRANSPORT THEORY*

GUNNAR ARONSSON† AND BENGT WINZELL‡

Abstract. In this paper we prove the existence of an equilibrium distribution in the three-dimensional diffusion model for suspended sediment transport. Mathematically, this means that we solve an elliptic boundary value problem with mixed boundary conditions in a domain with corners. We use Riesz–Schauder theory and some of the classical Agmon–Douglis–Nirenberg estimates, among other things.

Introduction. The problem of suspended sediment transport in a turbulent water stream has attracted very much attention, both theoretically and experimentally. One attempt to treat the problem mathematically is given by the so-called diffusion model with all its variations. See, for instance, [3, pp. 164–202 and pp. 398–419], or [9].

An equilibrium distribution is, roughly speaking, a distribution which is independent of the coordinate along the channel, independent of time, and satisfies certain boundary conditions. If one neglects the influence of the lateral coordinate, then the equilibrium distribution only depends on the vertical coordinate and is easily determined mathematically. This case has been known for a long time; see [3, pp. 172–173], and references given there.

If the influence of the lateral coordinate is *not* neglected, then the question of the equilibrium distribution leads to a boundary value problem for an elliptic differential equation in two variables under mixed boundary conditions. Furthermore, the boundary has corners. It is this problem that is treated in this paper.

We have proved that this boundary value problem has a unique solution. *Thus an equilibrium distribution exists and is unique.* Physically, we have in mind a waterflow or a channel with a fixed cross-section and a bottom which is not too steep. (A very steep bottom makes the boundary condition questionable.)

The paper is divided into a physical and a mathematical part, which can be read separately. The first part is mathematically elementary, whereas the second part requires knowledge of the theory of partial differential equations.

1. Physical assumptions for the diffusion model. We shall consider sediment transported in suspension by a turbulent water stream in an infinite channel. A number of simplifying assumptions must be made. To begin with, we consider the channel as being horizontal with a fixed cross-section Ω_0. We introduce a horizontal lateral coordinate x, a vertical coordinate y and a horizontal longitudinal coordinate z. The geometry of the problem is shown by Fig. 1.

We assume that the influence of the sediment on the water flow conditions can be neglected. This is reasonable, if the sediment concentration is not too high.

* Received by the editors November 26, 1975, and in revised form June 15, 1976.

† Department of Mathematics, Chalmers University of Technology and Göteborg University, S-402 20 Göteborg, Sweden.

‡ Department of Mathematics, Linköping University, S-581 83 Linköping, Sweden.

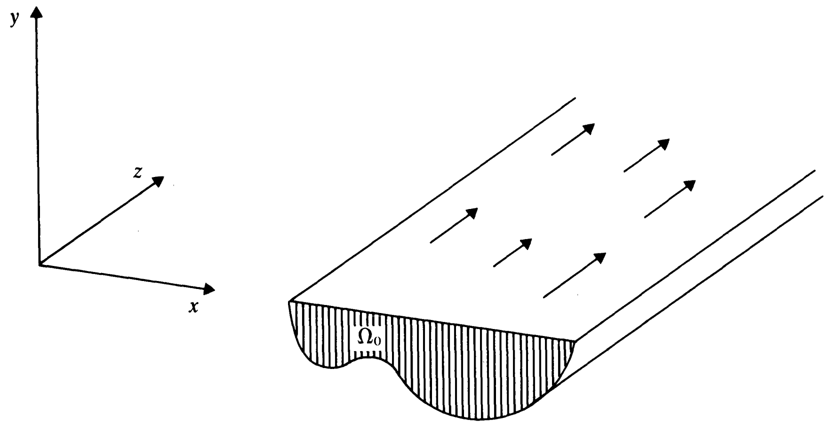

FIG. 1

We assume that the sediment has a well-defined settling velocity (the terminal fall velocity for a single particle in stagnant water).

We assume that the water flow conditions depend only on $(x, y) \in \Omega_0$. This means that the intensity of turbulence, and therefore the diffusion coefficient, and also the (mean) flow-velocity of the water, do not depend on the longitudinal coordinate z, or on time.

We introduce the following notations:

t = time,

$u(x, y, z, t)$ = concentration of sediment,

$\varphi(x, y)$ = diffusion coefficient,

w = settling velocity of sediment,

$\psi(x, y)$ = horizontal velocity of water, when turbulent variations have been "averaged out".

It is now part of our approach that the motion of sediment can be divided into the following three parts, which are superposed upon each other:

a) Isotropic turbulence of water gives a flux with direction

$$-(\partial u/\partial x, \partial u/\partial y, \partial u/\partial z)$$

and magnitude

$$\varphi(x, y) \cdot \sqrt{(\partial u/\partial x)^2 + (\partial u/\partial y)^2 + (\partial u/\partial z)^2};$$

b) the settling of particles causes a vertical flux with the magnitude $w \cdot u(x, y, z, t)$;

c) the horizontal mean velocity of water causes a flux in the positive z-direction with the magnitude $\psi(x, y) \cdot u(x, y, z, t)$.

2. The basic differential equation. By expressing mathematically the conservation of mass it follows from the above assumptions that the concentration of sediment u satisfies the parabolic equation

$$\frac{\partial u}{\partial t} = \varphi(x, y)\, \Delta u + u_x\varphi_x + u_y\varphi_y + wu_y - \psi u_z,$$

which is natural since it describes a diffusionlike process. Consequently, a stationary solution $u = u(x, y, z)$ must satisfy

$$\varphi(x, y)\,\Delta u + u_x\varphi_x + u_y\varphi_y + wu_y - \psi u_z = 0,$$

which is an elliptic equation. The reader will find in [3, Chap. 8] a more physically oriented presentation of this model, as well as further references.

3. Requirements on an equilibrium distribution. A boundary value problem. An equilibrium distribution of sediment is a stationary solution to our differential equation, which is independent of the longitudinal coordinate z, and which also satisfies certain boundary conditions. The boundary condition for the surface states the fact that the net vertical transport at the surface is zero, and the boundary condition for the bottom expresses a balance between the suspended load and the bottom load.

The net vertical flux is $wu + \varphi(x, y)\partial u/\partial y$, so the condition at the surface C_2 is simply $wu + \varphi(x, 0)\partial u/\partial y = 0$ there. See Fig. 2.

As for the bottom condition, we reason as follows. Consider an arbitrary vertical line L_ξ, where the depth is h. It is usually agreed *as a convention* that the transport at greater depth than $0.95h$ is dominated by the bottom transport. It is further assumed that, in a state of equilibrium, the suspended load must be adapted to the bottom load. Now we suppose that the bottom processes do not depend on z. It therefore seems reasonable to assume that the equilibrium distribution at $P \in C_1$ must be equal to a concentration $g(\xi)$, which is exactly the concentration that can be maintained by the bottom processes close to P. We therefore have, along C_1, a boundary condition of the form $u = g(x)$.

Summing up, we find that *the question of the existence of an equilibrium distribution leads to the following boundary value problem*:

Find a solution of the differential equation

$$\varphi(x, y)\,\Delta u + u_x\varphi_x + u_y\varphi_y + wu_y = 0$$

in the domain Ω, which satisfies the boundary conditions

$$u = g(x) \quad \text{on } C_1$$

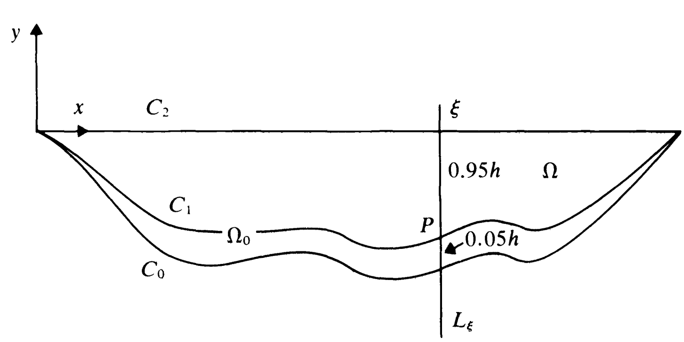

FIG. 2

and

$$wu + \varphi(x, 0)\frac{\partial u}{\partial y} = 0 \quad \text{on } C_2.$$

Here the function $\varphi(x, y)$ is positive and smooth in $\bar{\Omega}$, w is a positive constant and the function $g(x)$ is nonnegative and smooth. Further C_2 is a segment of the x-axis and C_1, the rest of $\partial\Omega$, is a smooth curve which can be represented by $y = h(x)$.

It will follow from the subsequent analysis that this boundary value problem has a unique solution and hence *there is a unique equilibrium distribution.*

The mathematical investigation which follows will require some further conditions on $\varphi(x, y)$, $h(x)$ and $g(x)$. But these conditions must be considered to be physically reasonable and will not be discussed further here.

4. Exact mathematical formulation of the boundary value problem. A uniqueness result. We are thus led to consider a problem of the form

$$\text{(P)} \qquad \begin{aligned} &\Delta u + a\frac{\partial u}{\partial x} + b\frac{\partial u}{\partial y} + cu = 0 \quad \text{in } \Omega, \\ &u = g \quad \text{on } \Gamma, \\ &\frac{\partial u}{\partial y} + \gamma u = 0 \quad \text{on } L^0. \end{aligned}$$

Here Ω is a bounded domain in the xy-plane, the boundary of which consists of the closed segment L on the x-axis and the curve Γ in the lower half-plane. Γ is given by $x \mapsto (x, h(x))$, $x \in L$, where $h \in C^{2+\alpha}(L)$. L^0 is the relative interior of L, and $h(x) < 0$ for $x \in L^0$.

For the coefficients of (P) we require

(i) $a, b \in C^{1+\alpha}(\bar{\Omega})$, $0 < \alpha < 1$,
(ii) $c \leqq 0$ belongs to $C^{\alpha}(\bar{\Omega})$,
(iii) $\gamma > 0$ belongs to $C^{2+\alpha}(L)$.

As a tool for uniqueness proofs we employ the Hopf maximum principles. See M. Protter and H. Weinberger [4, Thm. 6, p. 64, and Thm. 8, p. 67]. These theorems imply

LEMMA 4.1. *Let $u \in C(\bar{\Omega}) \cap C^1(\Omega \cup L^0) \cap C^2(\Omega)$ be a solution of* (P). *Then a nonnegative maximum (or a nonpositive minimum) is taken on at Γ. If in particular $g \geqq 0$, then $u \geqq 0$ in Ω and we get*

$$\max_{\bar{\Omega}} u = \max_{\Gamma} g.$$

COROLLARY. *A solution of the problem* (P) *is unique in the class $C(\bar{\Omega}) \cap C^1(\Omega \cup L^0) \cap C^2(\Omega)$.*

Remark. If $g > 0$ on Γ, it follows that $\min_{\bar{\Omega}} u > 0$, which seems physically satisfactory.

5. Transformations of the boundary value problem. We intend to derive existence results for (P) by a reflection method, similar to the classical Schwarz' reflection principle. Our present condition for u on L^0 is inconvenient for this, so we have to make a simple transformation into a problem with the condition $\partial u/\partial y = 0$ on L^0.

Having assumed that Γ is given by the graph of a function, it follows that the projection of $\bar{\Omega}$ on the x-axis coincides with L. Thus, by the formula

$$F(x, y) = y \cdot \gamma(x)$$

we have defined a function $F \in C^{2+\alpha}(\bar{\Omega})$ satisfying

$$\frac{\partial F}{\partial y}(x, 0) = \gamma(x) \quad \text{on } L. \tag{5.1}$$

LEMMA 5.1. *$u \in C(\bar{\Omega}) \cap C^1(L^0 \cup \Omega) \cap C^2(\Omega)$ is a solution of* (P) *if and only if the function*

$$v = e^F \cdot u, \tag{5.2}$$

having the same regularity, satisfies

$$\text{(P}'\text{)} \qquad \begin{aligned} &\Delta v + A\frac{\partial v}{\partial x} + B\frac{\partial v}{\partial y} + Cv = 0 \quad \text{in } \Omega, \\ &v = G \quad \text{on } \Gamma, \\ &\frac{\partial v}{\partial y} = 0 \quad \text{on } L^0, \end{aligned}$$

where

$$A = a - 2\frac{\partial F}{\partial x}, \quad B = b - 2\frac{\partial F}{\partial y}, \quad C = c + |\nabla F|^2 - \Delta F - a\frac{\partial F}{\partial x} - b\frac{\partial F}{\partial y} \quad \textit{and} \quad G = e^F g.$$

Remark. $A, B \in C^{1+\alpha}(\bar{\Omega})$, $C \in C^{\alpha}(\bar{\Omega})$.

COROLLARY. *A solution of the problem* (P′) *is unique within the class* $C(\bar{\Omega}) \cap C^1(\Omega \cup L^0) \cap C^2(\Omega)$.

The reflection technique. Denote by $\tilde{\Omega}$ the reflection of Ω in L, i.e.

$$\tilde{\Omega} = \{(x, -y) | (x, y) \in \Omega\}.$$

Put $D = \Omega \cup L^0 \cup \tilde{\Omega}$. Then D is a bounded domain in R^2. It is symmetric with respect to the x-axis. See Fig. 3. D has corners at the endpoints of L. The angles

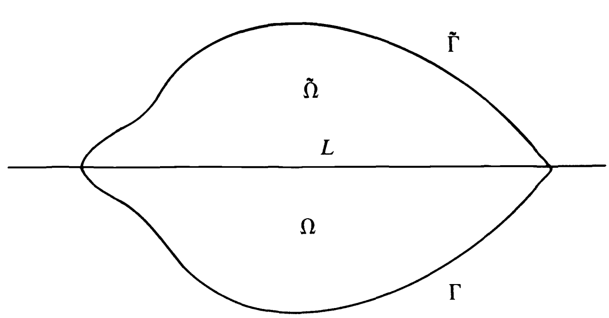

FIG. 3

are less than π. Denote by $\tilde{\Gamma}$ the reflection of Γ. Then $\partial D = \Gamma \cup \tilde{\Gamma}$. Define the symmetric extension of $\mathscr{L}' = \Delta + A\,\partial/\partial x + B\,\partial/\partial y + C$ to a differential operator in D by $\tilde{\mathscr{L}} = \Delta + \tilde{A}\,\partial/\partial x + \tilde{B}\,\partial/\partial y + \tilde{C}$ where $\tilde{A}$ and $\tilde{C}$ are even and $\tilde{B}$ is odd in y. We also extend G evenly in y to $\tilde{G}$.

The new problem to be investigated is a Dirichlet problem with a side condition:

$$
(\tilde{\mathrm{P}}) \qquad
\begin{aligned}
&\Delta \tilde{v} + \tilde{A}\frac{\partial \tilde{v}}{\partial x} + \tilde{B}\frac{\partial \tilde{v}}{\partial y} + \tilde{C}\tilde{v} = 0 \quad \text{in } D \backslash L,\\
&\tilde{v} = \tilde{G} \quad \text{on } \partial D,\\
&\tilde{v} \text{ is even in the variable } y,\\
&\tilde{v} \in C(\bar{D}) \cap C^1(D) \cap C^2(D\backslash L),
\end{aligned}
$$

where we assume that $\tilde{G}|_\Gamma = G \in C^{1+\alpha}(\Gamma)$.

Remark. $\tilde{A}$ is a Lipschitz function in $\bar{D}$ and restricted to $\bar{\Omega}$ or $\bar{\tilde{\Omega}}$ it is of class $C^{1+\alpha}$. $\tilde{C} \in C^\alpha(\bar{D})$ but $\tilde{B}$ is in general discontinuous at L. However, the restriction of $\tilde{B}$ to $\tilde{\Omega}$ has a continuous extension to $\bar{\tilde{\Omega}}$ which is in $C^{1+\alpha}$.

We immediately get

LEMMA 5.2. *A solution of* $(\tilde{\mathrm{P}})$ *is unique within* $C(\bar{D}) \cap C^1(D) \cap C^2(D\backslash L)$. *For given* G *on* Γ, *the problem* (P′) *has a solution* $v \in C(\bar{\Omega}) \cap C^1(\Omega \cup L^0) \cap C^2(\Omega)$ *if and only if the problem* $(\tilde{\mathrm{P}})$ *has an even in y solution* $\tilde{v} \in C(\bar{D}) \cap C^1(D) \cap C^2(D\backslash L)$ *with* $\tilde{v} = \tilde{G}$ *on* ∂D, *and in that case* $\tilde{v}$ *is the even (in y) extension of* v.

Remark. We may replace Ω by any domain with boundary consisting of one segment of the x-axis and a bounded curve in the lower half-plane.

6. Existence results for a modified problem. Since the domain in problem $(\tilde{\mathrm{P}})$ has corners we will approximate $(\tilde{\mathrm{P}})$ by problems in smooth domains.

In the sequel we will use the following convention: the functions a, b and c of problem (P) are extended (see [5, Th. 4, p. 177]) to all of R^2 as a', b' and c' with $c' \leqq 0$ and with preserved regularity. Then A', B' and C' are analogously defined in the lower half-plane. We get $\tilde{A}$, $\tilde{B}$ and $\tilde{C}$ on all of R^2 with symmetry and regularity as in § 5.

The following result is proved by standard arguments. (See also [10].)

LEMMA 6.1. *Assume that* Σ *is a bounded domain of* R^2, *symmetric with respect to the x-axis. Further, assume that* $\partial\Sigma$ *is of class* $C^{2+\alpha}$. *Let* $\mathscr{K}(x, y; \xi, \eta)$ *be the Green's function for the Laplacian in* Σ.

Let $w \in C^{2+\alpha}(\Sigma) \cap C^{1+\alpha}(\bar{\Sigma})$, *and let* $\tilde{G} \in C^{1+\alpha}(\partial\Sigma)$ *be even in y. Then w satisfies*

$$
(\mathrm{Q}) \qquad
\begin{aligned}
&\Delta w + \tilde{A}\frac{\partial w}{\partial x} + \tilde{B}\frac{\partial w}{\partial y} + \tilde{C}w = 0 \quad \text{in } \Sigma,\\
&w = \tilde{G} \quad \textit{on } \partial\Sigma,\\
&w \quad \textit{is even in } y
\end{aligned}
$$

if and only if

$$w(x, y) = -\int_{\partial\Sigma} \tilde{G}(\xi, \eta) \frac{\partial \mathcal{K}}{\partial \nu_{\xi,\eta}}(x, y; \xi, \eta)\, ds_{\xi,\eta} + \int_{\Sigma} \left(\tilde{A}\frac{\partial w}{\partial x} + \tilde{B}\frac{\partial w}{\partial y} + \tilde{C}w\right)(\xi, \eta)\mathcal{K}(x, y; \xi, \eta)\, d\xi\, d\eta \tag{6.1}$$

in Σ and w is even in y.

LEMMA 6.2. *Σ is as in the previous lemma. The integral equation* (6.1) has *a solution $w \in C^{2+\alpha}(\Sigma) \cap C^{1+\alpha}(\bar{\Sigma})$ which is even in y for every even (in y) $\tilde{G} \in C^{1+\alpha}(\partial\Sigma)$.*

Proof. Consider the linear operator

$$T: \begin{pmatrix} W \\ \tilde{G} \end{pmatrix} \mapsto \begin{pmatrix} w + \int_{\partial\Sigma} \tilde{G}\frac{\partial \mathcal{K}}{\partial \nu} - \int_{\Sigma}\left(\tilde{A}\frac{\partial w}{\partial x} + \tilde{B}\frac{\partial w}{\partial y} + \tilde{C}w\right)\mathcal{K} \\ \tilde{G} \end{pmatrix}$$

defined on $X = \{(w, \tilde{G}) \in C^1(\bar{\Sigma}) \times C^{1+\alpha}(\partial\Sigma)$: w and $\tilde{G}$ are even in $y\}$.

1) *T maps X into X.* The only fact that needs a proof is that the integral over the domain in the first component of $T\binom{w}{G}$ is in C^1. This, however, follows from an application of Theorem 9.3 of [1] and we have the result that in fact the map $f \mapsto \int_\Sigma f\mathcal{K}$ is continuous from $L^\infty(\Sigma)$ to $C^{1+\alpha}(\bar{\Sigma})$ for every $\alpha \in [0, 1[$. For the details we refer to our paper [10].

2) *$T - I = K$ is compact on X.* We already know that $K\begin{pmatrix} w \\ \tilde{G} \end{pmatrix} = \begin{pmatrix} v - u \\ 0 \end{pmatrix}$ with $v, u \in C^{1+\alpha}(\bar{\Sigma})$. Since $v \in C^{1+\alpha}(\bar{\Sigma})$ and $\Delta v = 0$ in Σ, Theorem 9.3 of [1] gives

$$\|v\|_{1+\alpha}^{\Sigma} \leqq C \cdot \|\tilde{G}\|_{1+\alpha}^{\partial\Sigma}.$$

and the argument of 1) shows that

$$\|u\|_{1+\alpha}^{\Sigma} \leqq C \cdot \|w\|_{1}^{\Sigma}.$$

Introduce the norm

$$\left\|\begin{pmatrix} w \\ \tilde{G} \end{pmatrix}\right\| = \|w\|_1^{\Sigma} + \|G\|_{1+\alpha}^{\partial\Sigma}$$

in X. We have shown that $T - I = K$ is compact by the Arzela–Ascoli theorem.

3) *The null space of T is trivial.* Assume that $(w, \tilde{G}) \in \mathcal{N}(T)$, the kernel of T. By definition of T it follows that $\tilde{G} = 0$ and that

$$w(x, y) = \int_{\Sigma}\left(\tilde{A}\frac{\partial w}{\partial x} + \tilde{B}\frac{\partial w}{\partial y} + \tilde{C}w\right)(\xi, \eta) \cdot \mathcal{K}(x, y; \xi, \eta)\, d\xi\, d\eta.$$

Since $w \in C^1(\bar{\Sigma})$ and $\tilde{A}, \tilde{B}, \tilde{C} \in L^\infty(\bar{\Sigma})$, it follows from [6] that $w \in C^{1+\alpha}(\bar{\Sigma})$. Furthermore, w is even in y, so $\partial w/\partial y(x, 0) = 0$. By the $C^{1+\alpha}$ regularity of the restrictions of $\tilde{B}$ to the components of $\Sigma\backslash\{x\text{-axis}\}$ it follows that $\tilde{A}\,\partial w/\partial x + \tilde{B}\,\partial w/\partial y + \tilde{C}w \in C^\alpha(\bar{\Sigma})$ and thus $\Delta w + \tilde{A}w + \tilde{B}w + \tilde{C}w = 0$ in Σ and $w \in C^{2+\alpha}(\bar{\Sigma})$. This is a minor modification of the result in [2, p. 250]. Since w has zero data it follows that w is identically zero.

4) *The range $R(T)$ of T is all of X according to the Riesz–Schauder theory* (see [8, Chap. X]). Thus the theorem is proved except for some regularity results which however are of standard type. Q.E.D.

7. An existence theorem for the basic boundary value problem. The notations Γ, Ω etc. now mean the same things as in § 4.

MAIN EXISTENCE THEOREM (Theorem 7.1). *For every $g \in C^{1+\alpha}(\Gamma)$ there is a unique solution $u \in C(\bar{\Omega}) \cap C^{1+\alpha}(\bar{\Omega} \backslash (\Gamma \cap L)) \cap C^2(\Omega)$ to the problem* (P).

Remark. By $u \in C^{1+\alpha}(\bar{\Omega} \backslash (\Gamma \cap L))$ it is meant that for every compact set K in $\bar{\Omega} \backslash (\Gamma \cap L)$, $u|_K \in C^{1+\alpha}(K)$.

Proof. The uniqueness was proved in § 4. It remains to prove the existence. The full details of this proof are given in [10] (available from the authors). The main difficulty is due to the fact that $D = \Omega \cup L^0 \cup \tilde{\Omega}$ is not of class $C^{2+\alpha}$. To circumvent this, we construct a sequence of $C^{2+\alpha}$-domains $D_n \searrow D$ and a corresponding sequence $\{V_n\}$ of solutions to problems similar to $(\tilde{\mathrm{P}})$ in D_n. Finally, we show that a subsequence of $\{V_n\}$ converges to a solution $\tilde{u}$ of $(\tilde{\mathrm{P}})$.

All that can be done by arguments based on a priori estimates from Theorem 9.3 of [1]. Such estimates also show that the limit function $\tilde{u}$ has the regularity stated in the theorem except possibly at the corners and that the function u which is the solution of (P) corresponding to $\tilde{u}$ satisfies the boundary conditions. To prove the continuity of u up to the corners we employ the theorem on bounded convergence and the estimates for Green's function in [6] and [7] to see that u satisfies the integral identity

$$\tilde{u}(x, y) = -\int_{\partial D} G\frac{\partial \mathcal{K}}{\partial \nu} + \int_D \tilde{u}\left\{-\frac{\partial}{\partial \xi}(\tilde{A}\mathcal{K}) - \tilde{B}\frac{\partial \mathcal{K}}{\partial \eta} + \tilde{C}\mathcal{K}\right\}$$
$$-\int_{\Omega} \tilde{u}\frac{\partial B}{\partial \eta}\mathcal{K} - \int_{\tilde{\Omega}} \tilde{u}\frac{\partial \tilde{B}}{\partial \eta}\mathcal{K} + 2\int_L uB\mathcal{K}.$$

Now the continuity follows in the same way as for smooth boundaries by the estimates in [6] and [7]. Q.E.D.

Remark. This main existence theorem can be generalized in various ways. For instance, we may allow Γ to have a finite number of *convex* corners. Widman's estimates for the Green's function still hold and our smoothing process can still be applied. Naturally, it is assumed that the sections of Γ connecting these corners are of class $C^{2+\alpha}$. We obtain a solution u to our boundary value problem in the class

$$U \in C^{2+\alpha}(\Omega) \cap C(\bar{\Omega}) \cap C^{1+\alpha}(\bar{\Omega}\backslash \text{all corners}).$$

Mathematically, this would enable us to include the case of an artificial channel with a convex, polygonal cross-section, for instance a rectangle. However, it is far from clear what the physical boundary conditions should be then, so it seems advisable to refrain from any physical statement in that case.

REFERENCES

[1] S. AGMON, A. DOUGLIS AND L. NIRENBERG, *Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions I*, Comm. Pure Appl. Math., 12 (1959), pp. 623–727.

[2] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, vol. II, Interscience, New York, 1962.

[3] W. H. Graf, *Hydraulics of Sediment Transport*, McGraw-Hill, New York, 1971.

[4] M. H. Protter and H. F. Weinberger, *Maximum Principles in Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1970.

[5] E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, NJ,1970.

[6] K. O. Widman, *Inequalities for the Green function and boundary continuity of the gradient of solutions of elliptic differential equations*, Math. Scand., 21 (1967), pp. 17–37.

[7] ———, *Inequalities for Green functions of second order elliptic operators*, Rep. 8—1972, Dept. of Math., Linköping Univ., Sweden, 1972.

[8] K. Yosida, *Functional Analysis*, Springer-Verlag, Berlin-Heidelberg-New York, 1968.

[9] W. W. Sayre, *Dispersion of silt particles in open channel flow*, Proc. Amer. Soc. Civil Engrs., 95 (1969), no. HY 3.

[10] G. Aronsson and B. Winzell, *Suspended sediment transport in open channel flow. Existence of an equilibrium distribution in the three-dimensional case.* Rep. LiTH-MAT-R-75-12, Dept. of Math., Linköping Univ., Sweden, 1975.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

THE ENERGY APPROACH TO THE KARMAN–FÖPPL EQUATIONS*

MELVIN MULLIN†

Abstract. It is shown that the energy of a nonlinear elastic plate, subject to an infinite set of constraints, attains its minimum. These constraints are equivalent to the stress-strain relations. It is also shown that the minimizing functions w, ϕ satisfy the Karman–Föppl equations:

$$N\Delta^2 w - \phi_{xx}w_{yy} - \phi_{yy}w_{xx} + 2\phi_{xy}w_{xy} = p,$$

$$\Delta^2\phi + E(w_{xx}w_{yy} - w_{xy}^2) = 0.$$

Introduction. An energy method is used to prove the existence of solutions to the nonlinear system of Karman–Föppl equations:

$$N\Delta^2 w - \phi_{xx}w_{yy} - \phi_{yy}w_{xx} + 2\phi_{xy}w_{xy} = p, \tag{1}$$

$$\Delta^2\phi + E(w_{xx}w_{yy} - w_{xy}^2) = 0. \tag{2}$$

The function w gives the deflection of an elastic plate occupying a plane region Ω and subject to a normal load $p(x, y)$; ϕ is the stress function; N and E are positive constants. The equation (1) may be formally derived as a necessary condition for the existence of minimum potential energy $E(w, \phi)$ subject to the constraining equation of compatibility (2). A basic difficulty in proving the existence of a minimizing pair (w, ϕ) is due to the fact that $E(w, \phi)$ contains only derivatives up to *second* order. This implies that the minimizing functions should be sought in a space of functions with square integrable second derivatives. The *fourth* order compatibility equation is not meaningful for such functions. In § 1 it will be shown how the constraint (2) can be replaced by an *infinite* set of integral constraints containing only second derivatives of ϕ and first derivatives of w. It will be shown in § 2 that if there is a smooth pair of functions (w, ϕ) that minimizes $E(w, \phi)$ over all pairs that satisfy the infinite set of constraints, then w and ϕ solve the Karman–Föppl equations (1), (2). The existence of such a pair will be established in § 3.

When no forces are applied at the boundary, the energy is given by

$$E(w, \phi) = \frac{N}{2}B(w) + \frac{1}{4}q_\phi(w) - (w, p), \tag{3}$$

where

$$B(w) = \int_\Omega\int (\Delta w)^2 - 2(1-\nu)(w_{xx}w_{yy} - w_{xy}^2), \tag{4}$$

$$q_\phi(w) = \int_\Omega\int \phi_{xx}w_y^2 + \phi_{yy}w_x^2 - 2\phi_{xy}w_xw_y, \tag{5}$$

$$(w, p) = \int_\Omega\int wp, \tag{6}$$

and the Poisson ratio ν satisfies $0 < \nu < 1$.

* Received by the editors December 11, 1975, and in revised form September 14, 1976.

† John Jay College of Criminal Justice, City University of New York, New York, New York. Now at Becker Securities Corporation, New York, New York 10041.

A problem arises in trying to find a lower bound for $q_\phi(w)$, the quadratic form in w defined by (5). If the plate is under tension at a point (x, y), then the integrand in $q_\phi(w)$ is positive there. It will be shown that $q_\phi(w)$ is positive *without* the a priori assumption that the plate is under tension.

Previous existence theorems have been based on the Schauder fixed point theorem (Knightly [7]), or have been established for small loads by perturbation techniques (Fife [4]). The energy approach presented here is somewhat more constructive than the use of fixed point theorems and is not restricted to small loads.

1. The set of constraints. The compatibility condition results from the elimination of the displacements u and v, in the x and y directions respectively, from the stress-strain relations:

$$u_x + \frac{1}{2} w_x^2 = \frac{1}{E}(\phi_{yy} - \nu\phi_{xx}), \tag{7}$$

$$v_y + \frac{1}{2} w_y^2 = \frac{1}{E}(\phi_{xx} - \nu\phi_{yy}), \tag{8}$$

$$u_y + v_x + w_x w_y = -\frac{2(1+\nu)}{E}\phi_{xy}. \tag{9}$$

This is accomplished by applying ∂_y^2 to (7), ∂_x^2 to (8), $-\partial_x\partial_y$ to (9) and adding. To avoid the fourth order equation (2), the following approach is taken here: Taking an arbitrary function $\psi \in C_0^\infty(\Omega)$, equation (7) is multiplied by ψ_{yy}, (8) by ψ_{xx}, and (9) by $-\psi_{xy}$. The sum is integrated over Ω, yielding

$$\frac{1}{2}\int_\Omega\int (\psi_{xx} w_y^2 + \psi_{yy} w_x^2 - 2\psi_{xy} w_x w_y) = \frac{1}{E}\int_\Omega\int \Delta\phi\,\Delta\psi, \quad \forall \psi \in C_0^\infty(\Omega),$$

or from the definition (5):

$$\frac{1}{2} q_\psi(w) = \frac{1}{E}\int_\Omega\int \Delta\phi\,\Delta\psi, \quad \forall\psi \in C_0^\infty(\Omega). \tag{10}$$

This is the weak form of the differential equation (2).

2. Necessary conditions. For simplicity we consider in detail only the case of a clamped plate to which no forces are applied at the boundary. It will be shown that if $(\bar{w}, \bar{\phi})$ is a pair of smooth functions that minimizes $E(w, \phi)$ over all functions (w, ϕ) that satisfy (10) and the boundary conditions

$$w = 0, \qquad w_n = 0 \qquad \text{on } \partial\Omega, \tag{11}$$

$$\phi = 0, \qquad \phi_n = 0 \qquad \text{on } \partial\Omega, \tag{12}$$

then $(\bar{w}, \bar{\phi})$ is a solution of the Karman–Föppl equations (1), (2).

Let $v \in C_0^\infty(\Omega)$ and set

$$w = \bar{w} + \varepsilon v. \tag{13}$$

Denote by $\phi(\varepsilon)$ the solution of the differential equation

$$\Delta^2\phi = -E(w_{xx}w_{yy} - w_{xy}^2), \tag{2}$$

subject to the boundary conditions (12). If we multiply (2) by any $\psi \in C_0^\infty(\Omega)$ and integrate by parts, we find that the pair $(w, \phi(\varepsilon))$ satisfies (10). Since the differential equation (2) has a unique solution ϕ satisfying (12) for a given w, we conclude that $\phi(0) = \bar{\phi}$ and that the pair $(\bar{w}, \bar{\phi})$ is a solution of (2). Furthermore, because $\phi(0) = \bar{\phi}$, $E(\bar{w} + \varepsilon v, \phi(\varepsilon))$ has a minimum at $\varepsilon = 0$ and therefore

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} E(\bar{w} + \varepsilon v, \phi(\varepsilon)) = 0. \tag{14}$$

From the definitions (3)–(6), this is equivalent to

$$NB(\bar{w}, v) + \tfrac{1}{2}q_{\bar{\phi}}(\bar{w}, v) + \tfrac{1}{4}q_{\phi_\varepsilon}(\bar{w}) - (x, p) = 0, \tag{15}$$

where

$$B(w, v) = \int_\Omega\int \Delta w\, \Delta v - (1-\nu)(w_{xx}v_{yy} + w_{yy}v_{xx} - 2w_{xy}v_{xy}), \tag{16}$$

$$q_\phi(w, v) = \int_\Omega\int \phi_{xx}v_y w_y + \phi_{yy}v_x w_x - \phi_{xy}(w_x v_y + w_y v_x), \tag{17}$$

$$\phi_\varepsilon = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \phi(\varepsilon). \tag{18}$$

We will show that

$$q_{\phi_\varepsilon}(\bar{w}) = 2q_{\bar{\phi}}(\bar{w}, v) \tag{19}$$

and therefore (15) may be written as

$$NB(\bar{w}, v) + q_{\bar{\phi}}(\bar{w}, v) - (v, p) = 0.$$

Integration by parts yields

$$\int_\Omega\int [N\,\Delta^2\bar{w} - \bar{\phi}_{xx}\bar{w}_{yy} - \bar{\phi}_{yy}\bar{w}_{xx} + 2\bar{\phi}_{xy}\bar{w}_{xy} - p]v = 0.$$

Since v is arbitrary, we obtain the equation (1):

$$N\,\Delta^2\bar{w} - \bar{\phi}_{yy}\bar{w}_{xx} - \bar{\phi}_{xx}\bar{w}_{yy} + 2\bar{\phi}_{xy}\bar{w}_{xy} = p.$$

It remains to prove (19). We differentiate the constraint (10) for $\phi = \phi(\varepsilon)$, $w = \bar{w} + \varepsilon v$ with respect to ε to obtain

$$\frac{2}{E}\int\int \Delta\bar{\phi}\,\Delta\phi_\varepsilon = \frac{1}{2}q_{\phi_\varepsilon}(\bar{w}) + q_{\bar{\phi}}(\bar{w}, v). \tag{20}$$

If we multiply (2) with $\varepsilon = 0$ by ϕ_ε and integrate by parts, we obtain

$$\int\int \Delta\phi_\varepsilon\,\Delta\bar{\phi} = \frac{E}{2}q_{\phi_\varepsilon}(\bar{w}). \tag{21}$$

Equations (20) and (21) yield the result (19).

3. Existence of a minimizing pair. We consider functions in the space $H^2(\Omega)$ and $H_0^2(\Omega)$. These are the completions of $C^\infty(\Omega)$ and $C_0^\infty(\Omega)$ in the norm:

$$\|u\|_2^2 = \int_\Omega \int u^2 + \int_\Omega \int (\nabla u)^2 + \int_\Omega \int (u_{xx}^2 + u_{yy}^2 + 2u_{xy}^2). \tag{22}$$

For functions w in $H_0^2(\Omega)$ we have

$$\frac{1}{K}\|w\|_2 \leqq B(w)^{1/2} \leqq K\|w\|_2. \tag{23}$$

THEOREM. *There is a pair of functions $(\bar{w}, \bar{\phi})$, with $\bar{w}$ and $\bar{\phi} \in H_0^2(\Omega)$, that minimizes $E(w, \phi)$ over all such pairs (w, ϕ) restricted by the infinite set of constraints* (10).

Proof. The set of pairs (w, ϕ), with $w, \phi \in H_0^2(\Omega)$, that satisfy (10) for all $\psi \in C_0^\infty(\Omega)$ will be denoted by $\mathscr{C}$. We observe that $\mathscr{C}$ is not empty; for if $w \in C_0^\infty(\Omega)$, then we may find a $\phi \in H_0^2(\Omega)$ by solving (2), (12). Multiplying (2) by $\psi \in C_0^\infty(\Omega)$ and integrating by parts, we obtain (10).

Since $(0, 0) \in \mathscr{C}$, we find that

$$m = \inf_{\mathscr{C}} E(w, \phi) \tag{24}$$

satisfies

$$m \leqq E(0, 0) = 0.$$

It follows that we may find a sequence $\{(w_j, \phi_j)\}$ in $\mathscr{C}$ such that

$$E(w_j, \phi_j) \leqq 0, \quad \forall j \tag{25}$$

and

$$\lim_{j\to\infty} E(w_j, \phi_j) = m. \tag{26}$$

We will show that the sequences $\{w_j\}$ and $\{\phi_j\}$ are bounded in $H_0^2(\Omega)$. Definition (3) of the energy together with inequality (25) yields

$$\frac{N}{2}B(w_j) + q_{\phi_j}(w_j) \leqq (w_j, p). \tag{27}$$

Since both sides of (10) are continuous linear functionals on $\psi \in H_0(\Omega)$, we may take $\psi = \phi_j$ to find

$$q_{\phi_j}(w_j) = \frac{2}{E}\int\int (\Delta\phi_j)^2 \geqq 0.$$

Thus

$$\frac{N}{2}B(w_j) \leqq (w_j, p) \leqq \|w_j\|_{L_2}\|p\|_{L_2}.$$

From the definition (22) of the norm in $H_0^2(\Omega)$ and inequality (23), we conclude that

$$\|w_j\|_2 \leqq \frac{2K^2}{N}\|p\|_{L_2}. \tag{28}$$

Thus the sequence $\{w_j\}$ is bounded in $H_0^2(\Omega)$.

It is well known [5] that

$$\frac{1}{K_1}\|\phi\|_2 \leqq \|\Delta\phi\|_{L_2} \leqq K_1\|\phi\|_2 \quad \text{for } \phi \in H_0^2(\Omega); \tag{29}$$

thus it suffices to show that $\{\Delta\phi_j\}$ is a bounded sequence in L^2. Equations (10) with $\psi = \phi_j$ imply

$$\|\Delta\phi_j\|_{L^2}^2 \leqq \frac{E}{2}\|\phi_j\|_2\|w_j\|_{1,4}^2 \leqq \frac{K_1E}{2}\|\Delta\phi_j\|_{L^2}\|w_j\|_{1,4}^2$$

or

$$\|\Delta\phi_j\|_{L^2} \leqq \frac{K_1E}{2}\|w_j\|_{1,4}^2, \tag{30}$$

where

$$\|u\|_{1,4}^4 = \iint u^4 + \iint u_x^4 + \iint u_y^4. \tag{31}$$

It can be shown [5] that

$$\|u\|_{1,4} \leqq K_2\|u\|_2 \quad \text{for } u \in H_0^2(\Omega). \tag{32}$$

Thus inequalities (28) and (30) yield

$$\|\Delta\phi_j\|_{L^2} \leqq K_3\|p\|_{L^2}^2,$$

i.e., $\{\Delta\phi_j\}$ is a bounded sequence in L^2.

Since a Hilbert space is weakly compact [2], we can find a subsequence $\{(w_{j_k}, \phi_{j_k})\}$ such that $\{w_{j_k}\}$ converges weakly in $H_0^2(\Omega)$ to a function $\bar{w}$, and $\{\phi_{j_k}\}$ converges weakly to $\bar{\phi}$ in $H_0^2(\Omega)$. We will show that $(\bar{w}, \bar{\phi}) \in \mathscr{C}$ and

$$E(\bar{w}, \bar{\phi}) \leqq m. \tag{33}$$

This means that $(\bar{w}, \bar{\phi})$ minimizes $E(w, \phi)$ on $\mathscr{C}$.

For convenience, we rename our subsequence $\{(w_j, \phi_j)\}$. Note that

$$\begin{aligned} |q_{\phi_j}(w_j) - q_{\bar{\phi}}(\bar{w})| &= |q_{\phi_j}(w_j - \bar{w}, w_j + \bar{w}) + q_{\phi_j - \bar{\phi}}(\bar{w})| \\ &\leqq K_4\|p\|_{L^2}^3\|w_j - \bar{w}\|_{1,4} + |q_{\phi_j - \bar{\phi}}(\bar{w})|. \end{aligned}$$

The first term on the right above tends to zero because $\{w_j\}$ converges to $\bar{w}$ weakly in $H_0^2(\Omega)$ and consequently in the norm (31) [5]. The second term tends to zero because $\{\phi_j\}$ converges to $\bar{\phi}$ weakly in $H_0^2(\Omega)$. Therefore

$$\lim_{j\to\infty} q_{\phi_j}(w_j) = q_{\bar{\phi}}(\bar{w}). \tag{34}$$

Since a norm is weakly lower semi-continuous [2],

(35)
$$\lim_{j\to\infty} B(w_j) \geqq B(\bar{w}).$$

It follows from (26), (34) and (35) that

$$m = \lim_{j\to\infty} E(w_j, \phi_j) \geqq E(\bar{w}, \bar{\phi})$$

It remains to show that $(\bar{w}, \bar{\phi}) \in \mathscr{C}$; i.e., $(\bar{w}, \bar{\phi})$ satisfies (10):

$$\frac{1}{2} q_\psi(\bar{w}) = \frac{1}{E}(\Delta\bar{\phi}, \Delta\psi), \quad \forall \psi \in C_0^\infty(\Omega).$$

This is true because the left side is a continuous function of w in the norm (31), the right side is a continuous function of ϕ in $H_0^2(\Omega)$, and $(w_j, \phi_j) \in \mathscr{C}$ for every j:

$$\frac{1}{2} q_\psi(\bar{w}) = \lim_{j\to\infty} \frac{1}{2} q_\psi(w_j) = \lim_{j\to\infty} \frac{1}{E}(\Delta\phi_j, \Delta\psi) = \frac{1}{E}(\Delta\bar{\phi}, \Delta\psi). \quad \text{Q.E.D.}$$

The minimizing pair $(\bar{w}, \bar{\phi})$ is a "weak solution" to the Karman–Föppl equations. It has been shown that such solutions are smooth [3]. In fact if the load p is Hölder continuous with exponent α, then $\bar{w} \in C^{4+\alpha}$ and $\bar{\phi} \in C^{6+\alpha}$. The proof is based upon the well known "boot strap" method [5] for proving regularity of weak solutions of elliptic partial differential equations together with results of Agmon [1] on L^p solutions to the Dirichlet problem.

Acknowledgment. The author would like to thank Fritz John for many interesting discussions.

REFERENCES

[1] S. Agmon, *The L_p Approach to the Dirichlet problem I*, Ann. Scuola Norm. Sup. Pisa, 13 (1959), pp. 405–448.

[2] N. Akhiezer and I. Glazman, *Theory of Linear Operators in Hilbert Space*, Ungar, New York, 1961–1963.

[3] M. Berger, *On von Karman's equations and the buckling of a thin elastic plate I*, Comm. Pure Appl. Math., 20 (1967), pp. 687–719.

[4] P. Fife, *Non-linear Deflection of Thin Elastic Plates under Tension*, Ibid., 14 (1961), pp. 81–112.

[5] A. Friedmann, *Partial Differential Equations*, 2nd ed., Krieger, New York, 1976.

[6] K. O. Friedrichs and J. J. Stoker, *The non-linear boundary value problem of the buckled plate*, Amer. J. Math., 63 (1944), pp. 839–888.

[7] G. Knightly, *An existence theorem for the von Karman equations*, Arch. Rational Mech. Anal., 27 (1967), pp. 233–242.

[8] S. Timoshenko, *Theory of Plates and Shells*, McGraw-Hill, New York, 1959.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

ON CONVEXITY PRESERVING OPERATORS*

RIVKA SENDEROVIZH†

Abstract. Let $K * f$ denote a convolution type operator, defined for periodic functions with period 2π; we present a characterization of the class of such operators preserving the set of odd functions which are concave on 2π. The paper concludes with some necessary conditions on the Fourier coefficients of functions of those classes.

1. Introduction. Let P be the class of odd functions $f(x)$ with period 2π which are L-integrable and nonnegative for $0 \leqq x \leqq \pi$.

Let C be the subclass of P consisting of all functions $f(x)$ of P which are concave in the interval $0 \leqq x \leqq \pi$.

Finally let the integral operator:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} K(x-y)f(y)\,dy = K * f(x) \tag{1.1}$$

where $K(x)$ is a real periodic function of bounded variation, be defined on P. The question is: under which condition does the operator (1.1) preserve the class C?

In [3], S. Karlin gives sufficient conditions for a kernel K which is twice continuously differentiable. He proves that if K is cyclic totally positive of order 3, and if all its Fourier coefficients are positive, then the suitable operator (1.1) preserves the class C and $K * f$ satisfies the inequality $K * f \leqq f$ on $[0, \pi]$.

In a way similar to that used by G. Pólya and I. J. Schoenberg in [4] for the de la Vallee–Poussin means, it can be proved that, if $K(x) = \sum_{-\infty}^{\infty} \mu_n e^{inx}$ is in SC_1 and SC_3 and has two continuous derivatives, and if $\mu_1 > 0$, then K solves problem.

In the sequel we use M. Fekete's [1] conclusions to find necessary and sufficient conditions for the above problem.

2. Necessary and sufficient conditions for a kernel to preserve the classes C and P.

DEFINITION 1. A periodic function is a *bell function* if it is nondecreasing on $(-\pi, 0)$ and nonincreasing on $(0, \pi)$.

LEMMA 1. *If $K(x)$ is an even bell function, then for each x and x_0 in $[0, \pi]$ the inequality $K(x-x_0) - K(x+x_0) \geqq 0$ holds.*

This property of an even bell function leads to the following theorem:

THEOREM 1. *Let $K(x)$ be a continuous even function of period 2π. Then, for each f of P, $K * f$ given by (1.1) belongs to P if and only if K is a bell function.*

Proof. If $K(x) = \sum_{-\infty}^{\infty} \mu_n e^{inx}$ is an even function, and $f(x) \sim \sum_{n=1}^{\infty} b_n \sin nx$, then

$$\begin{aligned} g(x) = K * f(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} K(x-y)f(y)\,dy \\ &= \frac{1}{2\pi}\int_{0}^{\pi} [K(x-y) - K(x+y)]f(y)\,dy \geqq 0 \quad \text{on } [0, \pi] \end{aligned}$$

* Received by the editors June 20, 1975, and in revised form July 12, 1976.

† Department of Mathematics, Technion—Israel Institute of Technology, Haifa, Israel. This paper is contained in an M.Sc. thesis submitted to the Technion, 1973.

whenever f belongs to P if and only if $[K(x-y)-K(x+y)]\geqq 0$ or equivalently $K(x)$ is a bell function.

Integration by parts yields

$$g^{(2n)}(x)=\frac{1}{2\pi}\int_{-\pi}^{\pi} K(x-y)f^{(2n)}(y)\,dy$$

which leads us to:

THEOREM 2. *Let f be an odd function satisfying $f^{(2n)}\geqq 0$ on $[0,\pi]$. Then $g(x)=K*f(x)$ satisfies $g^{(2n)}\geqq 0$ on $[0,\pi]$ if and only if K is an even bell function.*

For $n=1$ Theorem 2 gives us the condition under which K preserves the class C.

Remark 1. Using Theorems 1, 2 and the definition given by M. Fekete [1, p. 110] we find that a sequence of Fourier coefficients of an even bell function is one which preserves the classes P and C as well as the class of all odd functions satisfying $f^{(2n)}\geqq 0$ on $[0,\pi]$.

3. Some inequalities for the Fourier coefficients of a function of C and an even bell function. The class C is a cone, and its extreme rays are the family

$$f(a,b,x)=\begin{cases} \left.\begin{array}{ll} \dfrac{b}{a}x, & x\in[0,a], \\ b\dfrac{\pi-x}{\pi-a}, & x\in(a,\pi], \end{array}\right\} \text{ if } a\in(0,\pi); \\ \left.\begin{array}{ll} \dfrac{b}{\pi}(\pi-x), & x\in(0,\pi], \\ 0, & x=0, \end{array}\right\} \text{ if } a=0; \\ \left.\begin{array}{ll} \dfrac{b}{\pi}x, & x\in[0,\pi), \\ 0, & x=\pi, \end{array}\right\} \text{ if } a=\pi; \end{cases}$$

with $b>0$.

Using their Fourier series we get:

THEOREM 3. *If $f(x)\sim\sum_{n=1}^{\infty} b_n \sin nx$ belongs to C, then*:

(a) $\sum_{n=1}^{k} nb_n\geqq 0$, $k=1,2,\cdots$.

(b) $|b_n|\geqq k|b_{kn}|$, $n=1,2,\cdots$; $k=1,2,\cdots$.

Moreover, by Remark 1, we have

THEOREM 4. *If $K(x)=\sum_{-\infty}^{\infty}\mu_n e^{inx}$ is a continuous even bell function, then*:

(a) $\sum_{n=1}^{k}\mu_n\geqq 0$, $k=1,2,\cdots$.

(b) $|\mu_n|\geqq|\mu_{nk}|$, $n=1,2,\cdots$; $k=1,2,\cdots$.

Acknowledgment. The author wishes to express her deep gratitude to Professor Z. Ziegler for his guidance and inspiration.

REFERENCES

[1] M. FEKETE, *On certain classes of periodic functions and their Fourier series*, Bull. Res. Council of Israel, 7F (1958), pp. 103–112.

[2] A. JAKIMOVSKI, *Some remarks concerning "On certain classes of periodic functions and their Fourier series" of M. Fekete*, Ibid., 7F (1958), pp. 113–116.

[3] S. KARLIN, *Total Positivity and Applications*, vol. 1, Stanford University Press, Stanford, CA, 1968.

[4] G. PÓLYA AND I. J. SCHOENBERG, *Remarks on de la Vallée Poussin means and convex conformal maps of the circle*, Pacific J. Math., 8 (1858), pp. 295–334.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

SINGULAR PERTURBATION OF AN EXTERIOR DIRICHLET PROBLEM*

GEORGE C. HSIAO†

Abstract. This paper discusses a class of singular perturbation problems such as those of slow viscous flow past a cylinder. A semilinear second-order elliptic equation with a small parameter is used as a model to illustrate the correlation between a regular perturbation procedure and the method of matched asymptotic expansions. Some justification of the formal inner and outer expansions is established. It is found that the use of integral equations of the first kind for treating such a class of singular perturbation problems seems most desirable.

1. Introduction. In [8] we discuss the validity of the method of inner and outer expansions for treating singular perturbation problems such as those of slow viscous flow past a cylinder. The particular model we studied there is an ordinary differential equation problem:

$$(\hat{\mathrm{P}}_\varepsilon) \qquad \begin{gathered} y''+\frac{1}{x}y'-\varepsilon yy'=0 \quad \text{on } x>1, \\ y=0 \quad \text{at } x=1; \quad y\to -a \quad \text{as } x\to\infty, \end{gathered}$$

where ε is a small positive parameter and a is a positive constant independent of ε. Based on a regular perturbation procedure developed by Finn and Smith [4], [5], it is shown that the formal asymptotic expansions constructed by the method of inner and outer expansions are indeed in some sense the asymptotic expansions for the exact solution of the problem $(\hat{\mathrm{P}}_\varepsilon)$.

The purpose of this paper is to see how the ideas used for $(\hat{\mathrm{P}}_\varepsilon)$ might be extended and applied to similar problems for partial differential equations as a first step towards establishing the validity of the formal procedure for obtaining the inner and outer expansions in the case of full nonlinear Navier–Stokes equations. As a genuinely two-dimensional model,[1] we consider an exterior Dirichlet problem in the plane for the semilinear elliptic partial differential equation,

$$(\mathrm{E}) \qquad \Delta u-\varepsilon u u_{x_1}=0,$$

in an exterior domain Ω with a smooth boundary $\partial\Omega$ consisting of a simple closed curve. Here points in the plane E_2 are denoted by $\mathbf{x}=(x_1, x_2)$; $u=u(\mathbf{x};\varepsilon)$ and $u_{x_1}=\partial u/\partial x_1$. The boundary condition and condition at infinity are respectively,

$$(\mathrm{B}) \qquad u=f \quad \text{on } \partial\Omega$$

* Received by the editors September 24, 1974, and in final revised form August 17, 1976.

† Department of Mathematics, University of Delaware, Newark, Delaware 19711, and Fachbereich Mathematik, Technische Hochschule Darmstadt, D 61 Darmstadt, West Germany. This research was supported in part by the Alexander von Humboldt-Foundation, and in part by the U.S. Air Force Office of Scientific Research through Grant-AF-AFOSR 76-2879.

[1] Notice that $(\hat{\mathrm{P}}_\varepsilon)$ might be considered to represent a problem for the two-dimensional Laplace equation in an axially symmetric situation.

and

(C) $$u \to -a \quad \text{as} \; |\mathbf{x}| \to \infty.$$

For simplicity, we assume that the given function f is sufficiently smooth and independent of ε. Throuhgout the paper, we denote by (P_ε) the above problem.

Problems $(\hat{P}_\varepsilon)$ and (P_ε) belong to the special class of singular perturbation problems, a class whose prototype is the problem of the stationary incompressible flow of a viscous fluid past a cylinder [9], [11], [17]. Singular perturbation problems of this kind often arise in fluid mechanics, elasticity as well as in other fields of mathematical physics [10]. In this class of singular perturbation problems, the differential equations considered in general are of the form

$$L_\varepsilon u \equiv L_0 u + \varepsilon N u = 0, \tag{1.1}$$

where L_0 is a linear elliptic operator, and N is an operator which may or may not be linear but whose order is *less* than that of the operator L_0. The equation (1.1) is to hold in a region which is in some sense infinite. Solutions of (1.1) are to be subject to a boundary condition of the form (B) and a condition at infinity such as (C). The problems are *singular* in the sense that the degenerate problem (P_0),

$$L_0 u = 0 \tag{1.2}$$

together with the conditions (B) and (C), has no solution[2]. This is the analogue of the Stokes paradox in fluid flow [20], and thus, the degenerate equation, (1.2), is often referred to as the Stokes equation. In contrast to the usual singular perturbation problems considered in [23], [16], [3], neither the order nor the type of the degenerate equation, (1.2), has been changed from the original one, and the region of nonuniformity (or the boundary layer) in this case is the neighborhood of the point at infinity, rather than of the boundary [20, p. 153].

The degenerate problem (P_0), of course, will have solutions if the condition at infinity is relaxed. There will, in fact, be many solutions of (1.2) satisfying (B). More and more can be obtained by allowing increasingly singular behavior at infinity. Among them, there will be certain ones with the weakest possible singularities at infinity. We make use of what might be called the *weak singularity principle* (WSP) which states that only these weakest singular solutions should enter into the problem (compare [20, p. 53]). In problem (P_ε) we consider here, as one will see in § 2, WSP implies that the condition at infinity (C) should be replaced by the modified condition

(C′) $$u = A \log |\mathbf{x}| + O(1) \quad \text{as} \; |\mathbf{x}| \to \infty,$$

where A is any arbitrary constant. The question arises as to whether or not we can choose A so that the problem, $L_0 u = 0$ together with (B) and (C′), will give a meaningful result. It is here that one needs the *matching principle* in the singular perturbation theory. We will discuss this in § 3.

[2] Throughout this paper, unless otherwise specified, by a solution of the problem we always mean a solution in the classical sense. (e.g. u is a classical solution of (P_ε) if $u \in C^2(\Omega) \cap C(\bar{\Omega})$ and satisfies (E), (B) and (C)).

On the other hand, for the nonlinear operator N, there is a different kind of linearization which emphasizes the behavior at infinity. If we let $u = -a + v$, then v will satisfy an equation of the form

$$\mathscr{L}_\varepsilon v \equiv \mathscr{L} v + \varepsilon \mathscr{N} v. \tag{1.1'}$$

Here $\mathscr{L}$ is a linear operator depending on ε and hence different from L_0, while the operator $\mathscr{N}$ may or may not be the same as N. That there exists a unique solution to the linearized problem,

$$\mathscr{L} v = 0 \tag{1.2'}$$

together with conditions that $v = g$ on $\partial\Omega$ and v tends to zero at infinity, is the basis for the procedure of Finn and Smith [4]. Here g may be any given smooth function. In particular, one may set g equal to $a+f$ in view of the boundary condition (B). In fluid flow, (1.2′) is referred to as the Oseen equation [20], and the corresponding linearized problem as the Oseen problem. The solution of the full nonlinear problem is then sought as a regular perturbation of the solution of the Oseen problem (see § 5). This regular perturbation procedure, in fact, gives us a kind of asymptotic development for the solution of (P_ε) (see § 6).

For the nonlinear problem (P_ε), the main results can be summarized in the following two theorems.

THEOREM 1. *There exists a solution* $u(\mathbf{x}; \varepsilon)$ *of the problem* (P_ε) *defined by* (E), (B), C) *for* ε *sufficiently small.*

THEOREM 2. *Let* $\mathscr{D}$ *be any compact subset of* Ω *and let* $\mathscr{D}_\delta$ *denote the region* $\{\boldsymbol{\xi} \in E_2 : |\boldsymbol{\xi}| \geqq \delta\}$, *where* $\delta > 0$ *is a parameter. Then there exist functions* q_0, q_1 *defined for* $\mathbf{x} \in \mathscr{D}$, *and* Q_1 *defined for* $\boldsymbol{\xi} \neq \mathbf{0}$ *such that*

$$u(\mathbf{x}; \varepsilon) = q_0(\mathbf{x}) + q_1(\mathbf{x})(\log \varepsilon)^{-1} + O(\log \varepsilon)^{-2} \quad \text{as } \varepsilon \to 0^+$$

uniformly on $\mathscr{D}$, *and*

$$u\left(\frac{\boldsymbol{\xi}}{\varepsilon}; \varepsilon\right) = -a + Q_1(\boldsymbol{\xi})(\log \varepsilon)^{-1} + O(\log \varepsilon)^{-2} \quad \text{as } \varepsilon \to 0^+$$

uniformly on $\mathscr{D}_\delta$ *for any* $\delta > d$, *where* $d = \sup\{|\mathbf{x}| : \mathbf{x} \in \partial\Omega\}$. *Moreover, functions* q_0, q_1 *and* Q_1 *can be constructed by the matching principle.*

Remark. In view of standard results of singular perturbation theory [22], [13], and [21], one might expect that the solution of (P_ε) has a uniform representation of the form

(*) (a solution of $L_0 u = 0$) + (boundary layer terms) + (terms which tend to zero uniformly with ε).

Indeed, one will have the form (*) by constructing a composite expansion of the solution. This will be indicated in §§ 2 and 6.

We recall that the results for (P_ε) are similar to those for ($\hat{P}_\varepsilon$), although Theorem 2 here is not as complete as the one obtained in [8]. In recent years, there has been an increasing effort to apply the method of matched asymptotic expansions to Dirichlet problems for elliptic equations with small parameters. In surveying the literature, we see that either the degenerate equations are of lower order than the original ones or the domains under consideration are bounded. In sharp contrast to these, much less attention has been paid to the kind of problems considered here from the viewpoint of the rigorous justification of the formal procedure. A pertinent reference with respect to this aspect seems to be [6], where a fourth order ordinary differential equation problem has been used as a model to discuss the validity of the asymptotic matching principle developed there. As indicated in [8] for $(\hat{P}_\varepsilon)$, the matching principle in § 3 is a simplified version of the one in [6]. Some partial justification of formal procedures for the Lagerstrom model, a variational form of $(\hat{P}_\varepsilon)$, has been given in [2], [12] and recently in [18]. There are, of course, many papers concerning primarily the construction of the formal procedures in this connection. To mention a few, the special case of flow past a cylinder has been treated in detail in [11], [17], and [1]. It is our hope that the present investigation including [8] may shed some light on the validity of the method of matched asymptotic expansions for such a class of singular perturbation problems, in particular for problems concerning the viscous flow past obstacles.

The proof of Theorem 1 is given in § 5 and uses estimates for solutions of linear problems. These estimates are obtained in § 4. Theorem 2 is established in § 6 based on the asymptotic development for the linear problem (the Oseen problem) in § 2 and the matching principle formulated in § 3.

2. The linear problems. In this section, we would like to devote our attention to the linear problem (the Oseen problem for our model (P_ε))

$$(P_0') \qquad \begin{aligned} \mathscr{L}v \equiv \Delta v + \alpha v_{x_1} &= 0 \quad \text{in } \Omega, \\ v &= g \quad \text{on } \partial\Omega, \\ v &\to 0 \quad \text{as } |\mathbf{x}| \to \infty, \end{aligned}$$

where $\alpha \equiv \varepsilon\alpha$ and g is a given smooth function.[3] We note that the linear problem (P_0') is singular according to our definition, if $\int_{\partial\Omega} g(\mathbf{x})\, dS_{\mathbf{x}} \neq 0$. Hence in order to gain some insight of the singular nature of the nonlinear problem (P_ε), it is natural to begin with a study of the asymptotic behavior of the solution of (P_0'). Our approach here is based on a method of integral equations of the first kind developed in [9]. It is found that the use of single layer potential for treating singular perturbation problems of this kind is particularly desirable from the viewpoint of constructing asymptotic expansions.

In this connection, we also consider the modified degenerate problem

$$(P_0'') \qquad \begin{aligned} \Delta u &= 0 && \text{in } \Omega, \\ u &= g && \text{on } \partial\Omega, \\ u &= A \log|\mathbf{x}| + O(1) && \text{as } |\mathbf{x}| \to \infty \end{aligned}$$

for a given constant A. It was mentioned in the Introduction that the solution of (P_0'') does exist and is unique for every fixed A. We will explore this idea here again by the use of single layer potential (see [9]).

[3] In what follows, it suffices to assume that g is of class $C^{1+\lambda}(\partial\Omega)$, the class of Hölder continuously differentiable functions on $\partial\Omega$ with exponent $0<\lambda<1$.

In what follows we write $v(x;\alpha;g)$ for the solution of (P_0') (which exists as we will see), and write $u_A(\mathbf{x};g)$ for the unique solution of (P_0''). We recall that the fundamental singularity of the equation $\mathscr{L}v=0$ is

$$(2.1)\qquad S(\mathbf{x}-\mathbf{y};\alpha)=-\frac{1}{2\pi}e^{-(\alpha/2)(x_1-y_1)}K_0\Big(\frac{\alpha}{2}|\mathbf{x}-\mathbf{y}|\Big).$$

Here $K_0(r)$ is the zeroth order modified Bessel function of the second kind which decays to zero as $r\to\infty$ in contrast to the logarithmic singularity of the Laplace equation. For $\mathbf{x}, \mathbf{y}$ in a compact set, the series development of K_0 takes the form

$$(2.2)_1\qquad K_0\Big(\frac{\alpha}{2}|\mathbf{x}-\mathbf{y}|\Big)=-\log|\mathbf{x}-\mathbf{y}|+(\Gamma_0-\log\alpha)+L(\mathbf{x};\mathbf{y};\alpha)$$

where Γ_0 is a constant and the series $L(\mathbf{x};\mathbf{y};\alpha)$,

$$(2.2)_2\qquad L(\mathbf{x};\mathbf{y};\alpha)=\sum_{k=1}^{\infty}\{a_k(\mathbf{x},\mathbf{y})\log\alpha+b_k(\mathbf{x},\mathbf{y})\}\alpha^{2k},$$

converges uniformly on compact subsets. Here

$$a_k(\mathbf{x},\mathbf{y})=\Gamma_k|\mathbf{x}-\mathbf{y}|^{2k},\qquad b_k(\mathbf{x},\mathbf{y})=(\gamma_k+\Gamma_k\log|\mathbf{x}-\mathbf{y}|)|\mathbf{x}-\mathbf{y}|^{2k}$$

with constants Γ_k and γ_k.

We now formulate the fundamental theorem on the asymptotic representation of $v(\mathbf{v};\alpha;g)$:

THEOREM 3. *Let $\mathscr{D}$ be a compact subset of Ω and let $\mathscr{D}_{\delta_0}$ denote the region $\{x\in E_2: |\mathbf{x}|\geqq\delta_0\}$, where $\delta_0>0$ is a parameter. Then we have*

$$(2.3)\quad v(\mathbf{x};\alpha;g)=u_0(\mathbf{x};g)+u_{-m}(\mathbf{x};0)(\log\alpha)^{-1}+O(\log\alpha)^{-2}\quad \textit{as } \alpha\to 0^+,$$

uniformly on $\mathscr{D}$; and

$$(2.4)\quad v(\mathbf{x};\alpha;g)=me^{-(\alpha/2)x_1}K_0\Big(\frac{\alpha}{2}|\mathbf{x}|\Big)(\log\alpha)^{-1}+O(\log\alpha)^{-2}\quad \textit{as } \alpha\to 0^+,$$

uniformly on $\mathscr{D}_{\delta_0}$ for any $\delta_0>d/\alpha$, where $d=\sup\{|\mathbf{x}|: x\in\partial\Omega\}$ and m is a linear functional of g.

Remark. Expansions (2.3) and (2.4) are usually referred to respectively as the *inner* and *outer* expansions, which will be discussed in § 3.

This theorem needs some explanation. It yields a kind of asymptotic expansion for the solution but reflects the nonuniformity of the expansion. One can form a composite expansion from (2.3) and (2.4), which is uniformly valid for all $\mathbf{x}\in\bar{\Omega}$. This can be done by introducing a mollifier ψ_ρ, an infinitely differentiable function of $\mathbf{x}\in E_2$ defined by

$$(2.5)\qquad \psi_\rho(\mathbf{x})=\begin{cases}1 & \text{for } |\mathbf{x}|\geqq r+\dfrac{\rho}{2},\\[2ex] 0 & \text{for } |\mathbf{x}|\leqq r-\dfrac{\rho}{2},\end{cases}$$

where ρ is a small positive number such that $0<\rho<r-d$, and $2r$ is the diameter of a compact set containing $\bar{\Omega}^C=E_2\backslash\Omega$. Then, the solution $v(\mathbf{x};\alpha;g)$ of problem (P_0')

has the uniform asymptotic representation:

$$v(\mathbf{x};\alpha;g)=w(\mathbf{x};\alpha)+v_\alpha(\alpha\mathbf{x};\alpha)+z(\mathbf{x};\alpha) \tag{2.6}$$

where

$$w(\mathbf{x};\alpha)=(1-\psi_\rho(\alpha\mathbf{x}))\{u_0(\mathbf{x};g)+u_{-m}(\mathbf{x};0)(\log\alpha)^{-1}\},$$

$$v_\alpha(\alpha\mathbf{x};\alpha)=\psi_\rho(\alpha\mathbf{x})\left\{me^{-(\alpha/2)x_1}K_0\left(\frac{\alpha}{2}|\mathbf{x}|\right)(\log\alpha)^{-1}\right\};$$

$$z(\mathbf{x};\alpha)=O(\log\alpha)^{-2}\quad \text{as } \alpha\to 0^+$$

uniformly on $\bar{\Omega}$. Note that in this case, the boundary layer is the neighborhood of the point at infinity. Hence the term $v_\alpha(\alpha(\mathbf{x};\alpha)$ in (2.6) corresponds to the boundary layer term according to [21].

The proof of Theorem 3 is based on the following two lemmas:

LEMMA 2.1. *For given A and g, the provlem* (P_0'') *has a unique solution in the form*:

$$u_A(\mathbf{x};g)=U^0(\sigma;\mathbf{x})-m_g, \tag{2.7}$$

where m_g is a fixed constant depending on g; σ satisfies the integral equation, $U^0(\sigma;\mathbf{x})=g(\mathbf{x})+m_g$, $\mathbf{x}\in\partial\Omega$, and $U^0(\sigma;\mathbf{x})$ is defined by

$$U^0(\sigma;\mathbf{x})\equiv\int_{\partial\Omega}\sigma(\mathbf{y})\log|\mathbf{x}-\mathbf{y}|\,dS_{\mathbf{y}}. \tag{2.8}$$

LEMMA 2.2. *For given g, the problem* (P_0') *has a unique solution in the form*:

$$v(\mathbf{x};\alpha;g)=U(\phi;x;\alpha), \tag{2.9}$$

where ϕ satisfies the integral solution, $U(\phi;\mathbf{x};\alpha)=g(\mathbf{x})$, $\mathbf{x}\in\partial\Omega$, and $U(\phi;\mathbf{x};\alpha)$ is defined by

$$U(\phi;\mathbf{x};\alpha)\equiv-e^{-(\alpha/2)x_1}\int_{\partial\Omega}\phi(\mathbf{y})K_0\left(\frac{\alpha}{2}|\mathbf{x}-\mathbf{y}|\right)dS_{\mathbf{y}}. \tag{2.10}$$

Remark. The function $e^{(\alpha/2)x_1}U$ is a solution of the equation $\Delta w-(\alpha^2/4)w=0$ for any smooth function ϕ (e.g. $\phi\in C^\lambda(\partial\Omega)$).

A proof of Lemma 2.1 is essentially contained in [9]. We will, however, repeat the proof here so that we have enough information about the density function σ to see how the asymptotic developments (2.3) and (2.4) are derived. The proof of Lemma 2.2 will be omitted, since the proof is similar to that of Lemma 2.1., and the property of the density function ϕ can be obtained from that of σ.

Proof of Lemma 2.1. We begin the proof by seeking a solution of the problem (P_0'') in the form (2.8). For any σ continuous on $\partial\Omega$, $U^0(\sigma;\mathbf{x})$ is harmonic in Ω. We determine σ by requiring that

$$U^0(\sigma;\mathbf{x})=g(\mathbf{x})+m_g,\qquad \mathbf{x}\in\partial\Omega. \tag{2.11}$$

Differentiating (2.11) with respect to the arc length $S_{\mathbf{x}}$ along $\partial\Omega$, we obtain the integral equation

$$\dot{g}(\mathbf{x}) = \int_{\partial\Omega} \sigma(\mathbf{y}) \frac{\partial}{\partial S_{\mathbf{x}}} \log|\mathbf{x}-\mathbf{y}|\, dS_{\mathbf{y}} \equiv M(\sigma). \tag{2.12}$$

Here the dot indicates differentiation with respect to $S_{\mathbf{x}}$. Equation (2.12) is a singular integral equation of first kind. For the theory of such equations see [15]. The homogeneous equation adjoint to (2.12) can be shown to have the unique linearly independent solution $\tilde{\sigma} \equiv 1$. The condition of solvability of (2.12) is hence fulfilled. The general solution of (2.12) has the form

$$\sigma(\mathbf{y}) = M^{-1}(\dot{g}) + \beta_g \Phi_0(\mathbf{y}), \tag{2.13}$$

where $\Phi_0(\mathbf{y})$, $\int_{\partial\Omega} \Phi_0(\mathbf{y})\, dS_{\mathbf{y}} \neq 0$, is a fixed, nontrivial solution of the homogeneous equation corresponding to (2.12), β_g is an arbitrary constant, and $M^{-1}(\dot{g})$ has the resolvent form

$$M^{-1}(\dot{g}) = \dot{g}(\mathbf{y}) + \int_{\partial\Omega} \dot{g}(\mathbf{x}) T(\mathbf{y}; \mathbf{x})\, dS_{\mathbf{x}} \tag{2.14}$$

with a continuous resolvent T. We choose β_g so that

$$\int_{\partial\Omega} \sigma(\mathbf{y})\, dS_{\mathbf{y}} = A, \tag{2.15}$$

and set

$$\hat{u}_A(\mathbf{x}; g) = U^0(\sigma; \mathbf{x}) = U^0(M^{-1}(\dot{g}); \mathbf{x}) + \beta_g U^0(\Phi_0; \mathbf{x}). \tag{2.16}$$

Then $\hat{u}_A(\mathbf{x}, g)$ is harmonic in Ω and equal to $g(\mathbf{x}) + m_g$ on $\partial\Omega$, where m_g is a constant defined by

$$m_g = U^0(M^{-1}(\dot{g}); \mathbf{x}) - g(\mathbf{x}), \qquad \mathbf{x} \in \partial\Omega. \tag{2.17}$$

Here we have used the fact that

$$U^0(\Phi_0; \mathbf{x}) \equiv 0 \quad \text{on } \partial\Omega, \tag{2.18}$$

which can be proved by an argument similar to that in [9]. Furthermore, we have

$$\hat{u}_A(\mathbf{x}; g) = A \log|\mathbf{x}| + O\left(\frac{1}{|\mathbf{x}|}\right) \quad \text{as } |\mathbf{x}| \to \infty.$$

Hence

$$u_A(\mathbf{x}; g) = \hat{u}_A(\mathbf{x}; g) - m_g \tag{2.19}$$

is the unique solution of (P_0'').

We return now to the proof of Theorem 3. From Lemma 2.2., the solution of (P_0') has the form (2.10),

$$v(\mathbf{x}; \alpha; g) \equiv U(\phi; \mathbf{x}; \alpha),$$

where ϕ satisfies the integral equation of the first kind,

$$U(\phi; \mathbf{x}; \alpha) = g(\mathbf{x}), \qquad \mathbf{x} \in \partial\Omega. \tag{2.20}$$

By the series development of K_0 in (2.2), equation (2.20) can then be written in the form,

$$(2.21)\quad \begin{aligned} g(\mathbf{x})+\sum_{n=1}^{\infty} g_n(\mathbf{x})\alpha^n &= \int_{\partial\Omega} \phi(y)\log|\mathbf{x}-\mathbf{y}|\,dS_{\mathbf{y}}-(\Gamma_0-\log\alpha)\int_{\partial\Omega}\phi(\mathbf{y})\,dS_{\mathbf{y}} \\ &\quad -\int_{\partial\Omega} L(\mathbf{x}-\mathbf{y};\alpha)\phi(\mathbf{y})\,dS_{\mathbf{y}} \end{aligned}$$

with $g_n(\mathbf{x})=(1/n!)(x_1/2)^n g(\mathbf{x})$. Differentiating (2.21) with respect to arc length $S_{\mathbf{x}}$ along $\partial\Omega$ and rearranging terms, we obtain

$$(2.22)\qquad \dot{g}+\sum_{n=1}^{\infty}\dot{g}_n\alpha^n+\frac{\partial}{\partial S_{\mathbf{x}}}\int_{\partial\Omega} L(\mathbf{x};\mathbf{y};\alpha)\phi(\mathbf{y})\,dS_{\mathbf{y}}=M(\phi),$$

with $M(\phi)$ as in (2.12). Since the left hand side of (2.22) is orthogonal (in the L_2 sense) to solutions (that is, constants) of the homogeneous adjoint equation, we can invert it. The inverse can be written as a resolvent term plus an arbitrary constant times Φ_0 as in (2.13). When the resolvent is applied to the terms in (2.22) involving $L(\mathbf{x};\mathbf{y};\alpha)$ and the order of integration is interchanged, we obtain an equation of the form

$$(2.23)\qquad \phi=M^{-1}(\dot{g})+\sum_{n=1}^{\infty}M^{-1}(\dot{g}_n)\alpha^n+\beta\Phi_0+\int_{\partial\Omega}K(\mathbf{x};\mathbf{y};\alpha)\phi(\mathbf{y})\,dS_{\mathbf{y}},$$

where the linear transformation $K(\mathbf{x};\mathbf{y};\alpha)$ is $O(\alpha^2\log\alpha)$ and β is some constant to be determined.

We can solve (2.23), by successive approximations, in the form

$$(2.24)\qquad \phi=M^{-1}(\dot{g})+\beta_{\Phi_0}+\chi,$$

where $\chi=O(\alpha)$. We set $\beta=\beta_g+\beta_0$, where β_g is defined by (2.15) with $A=0$, and β_0 is to be determined. Then,

$$(2.25)\qquad \int_{\partial\Omega}\phi(\mathbf{y})\,dS_{\mathbf{y}}=\beta_0\int_{\partial\Omega}\Phi_0(\mathbf{y})\,dS_{\mathbf{y}}+O(\alpha)$$

and

$$(2.26)\qquad \int_{\partial\Omega}\phi(\mathbf{y})\log|\mathbf{x}-\mathbf{y}|\,dS_{\mathbf{y}}=\hat{u}_0(\mathbf{x};g)+\beta_0U^0(\Phi_0;\mathbf{x})+O(\alpha)$$

(cf. (2.8) and (2.16)). By making use of (2.17), (2.18), (2.24), and (2.25), we obtain

$$(2.27)\qquad \beta_0=\frac{m_g}{(\int_{\partial\Omega}\Phi_0(\mathbf{y})\,dS_{\mathbf{y}})(\Gamma_0-\log\alpha)}+O(\alpha).$$

Consequently, for $\mathbf{x}$ in a compact subset $\mathscr{D}$ of Ω, it follows easily from (2.25)–(2.27) that if one uses the series $(2.2)_1$, then

$$(2.28)\quad \begin{aligned} v(\mathbf{x};\alpha;g)&=\hat{u}_0(\mathbf{x};g)-m_g-\frac{m_g}{\int_{\partial\Omega}\Phi_0(\mathbf{y})\,dS_{\mathbf{y}}}\left\{\frac{1}{\log\alpha}+\frac{\Gamma_0}{(\log\alpha)^2}+O\left(\frac{1}{\log\alpha^3}\right)\right\} \\ &\quad\cdot U^0(\Phi_0;\mathbf{x})+O(\alpha\log\alpha)\quad \text{as } \alpha\to 0^+, \end{aligned}$$

uniformly on $\mathscr{D}$. Now let

$$m = m_g. \tag{2.29}$$

Then if we compare (2.28) with (2.19) ($A = 0$), we see that the first two terms are just $u_0(\mathbf{x}; g)$ and the coefficient of $1/\log \alpha$ is precisely $u_{-m}(\mathbf{x}; 0)$. Moreover, $|m_g| \leqq \text{const.} \|g\|_1$, where

$$\|g\|_n = \max_{\mathbf{x}\in\partial\Omega} \left|\frac{d^k g}{dS^k}\right|, \qquad 0 \leqq k \leqq n,$$

and S is the arc length. This proves the first part of Theorem 3.

To prove (2.4), we see that the representation (2.10) implies

$$v(\mathbf{x}; \alpha; g) = -\left(\int_{\partial\Omega} \phi(\mathbf{y})\, dS_{\mathbf{y}}\right) e^{-(\alpha/2)x_1} K_0\left(\frac{\alpha}{2}|\mathbf{x}|\right) + \mathscr{R}(\mathbf{x}; \alpha), \tag{2.30}$$

where $\mathscr{R}(\mathbf{x}; \alpha) = O(\alpha)$ as $\alpha \to 0^+$ uniformly for $|\mathbf{x}| > d/\alpha$, $d = \sup\{|\mathbf{x}| : x \in \partial\Omega\}$. Then the result (2.4) follows easily from (2.25), (2.27), and (2.29). This completes the proof of Theorem 3.

3. Inner and outer expansions. With the help of the preliminary analysis in § 2, we now propose a formal procedure for obtaining what are usually called the *inner* and *outer* expansions of the solution to the nonlinear problem (P_ε). This procedure is based on a matching principle similar to the one used in [8]. To illustrate the idea, we shall describe the procedure (matching principle) by computing the first few terms of the former inner and outer expansions. Then, we show that this procedure can, in principle, be continued to be used for obtaining higher order terms. We believe that it will be true this actually yields an asymptotic expansion for the exact solution of (P_ε), although we can only carry out the verification up to the term of order $(\log \varepsilon)^{-2}$ (see § 6).

We begin with the formal inner expansion,

$$u(\mathbf{x}; \varepsilon) \sim u_0(\mathbf{x}; f) + \sum_{k=1}^{\infty} \frac{u_{A_k}(\mathbf{x}; 0)}{(\log \varepsilon)^k}. \tag{3.1}$$

Here $u_0(\mathbf{x}; f)$ and $u_{A_k}(\mathbf{x}; 0)$ are solutions of the problem (P_0'') with A, g replaced by 0, f and A_k, 0, respectively; that is, these are solutions of the Laplace equation subject to the conditions:

$$\begin{aligned} &u_0(\mathbf{x}; f) = f \quad \text{on } \partial\Omega, \qquad &&u_0(\mathbf{x}; f) = O(1) \quad \text{as } |x| \to \infty, \\ &u_{A_k}(x; 0) = 0 \quad \text{on } \partial\Omega, \qquad &&u_{A_k}(\mathbf{x}; 0) = A_k \log|\mathbf{x}| + O(1) \quad \text{as } |\mathbf{x}| \to \infty. \end{aligned} \tag{3.2}$$

The A_k's, $k \geqq 1$, are constants to be determined by the matching principle which will be stated later.

Remark. To be more precise, in general one requires the inner expansion to satisfy the equation (E) and the condition (B) but not (C). Instead, the condition (C) is replaced by the condition (C′). It is here that one needs the matching principle to choose a unique constant A. Thus, we refer to (C′) as the *matching condition.*

Despite the undetermined constants A_k, $k \geqq 1$, we have here the representations

$$u_0(\mathbf{x}; f) = \int_{\partial\Omega} \sigma(\mathbf{y}) \log |\mathbf{x}-\mathbf{y}|\, dS_{\mathbf{y}} - m_f; \qquad \sigma(\mathbf{y}) = M^{-1}(\dot{f}) + \beta_f \Phi_0 \tag{3.3}$$

and

$$u_{A_k}(\mathbf{x}; 0) = \frac{A_k}{\int_{\partial\Omega} \Phi_0(\mathbf{y})\, dS_{\mathbf{y}}} \int_{\partial\Omega} \Phi_0(\mathbf{y}) \log |\mathbf{x}-\mathbf{y}|\, dS_{\mathbf{y}}, \qquad k \geqq 1. \tag{3.4}$$

The functions Φ_0, $M^{-1}(\dot{f})$ and the constant β_f are defined in the same way as in § 2. We note that from (3.3) and (3.4) the matching conditions read

$$\begin{aligned} u_0(\mathbf{x}; f) &= -m_f + O\left(\frac{1}{|\mathbf{x}|}\right) \quad \text{as } |\mathbf{x}| \to \infty \\ u_{A_k}(\mathbf{x}; f) &= A_k \log |\mathbf{x}| + O\left(\frac{1}{|\mathbf{x}|}\right) \quad \text{as } |\mathbf{x}| \to \infty. \end{aligned} \tag{3.5}$$

Next we shall construct the outer expansion. To this end, we introduce the outer variable[4] $\boldsymbol{\xi} = \varepsilon \mathbf{x}$ and set $U(\boldsymbol{\xi}; \varepsilon) = u(\boldsymbol{\xi}/\varepsilon; \varepsilon)$. We denote the domain Ω by $\Omega_{\boldsymbol{\xi}}$ and adapt the similar convention for other notations in connection with the outer variable.

Observe that, in terms of the outer variable, the equation (E) becomes

$$\Delta_{\boldsymbol{\xi}} U = U U_{\xi_1} \quad \text{in } \Omega_{\boldsymbol{\xi}} \tag{E$'$}$$

and conditions (B) and (C) become, respectively

$$U = F(\boldsymbol{\xi}) \quad \text{on } \partial\Omega_{\boldsymbol{\xi}} \tag{B$'$}$$

and

$$U \to -a \quad \text{as } |\boldsymbol{\xi}| \to \infty. \tag{C$''$}$$

The outer expansion is of the form

$$U(\boldsymbol{\xi}; \varepsilon) \sim -a + \sum_{k=1}^{\infty} \frac{U_k(\boldsymbol{\xi})}{(\log \varepsilon)^k}. \tag{3.6}$$

This expansion is required to satisfy the equation (E′) and the condition (C″) *but not* (B′). Formally substituting (3.6) into (E′), (C″) and equating coefficients of like powers of $(\log \varepsilon)^{-1}$, one obtains the conditions for the functions $U_k(\boldsymbol{\xi})$; that is,

$$\begin{aligned} \mathscr{L} U_k \equiv \Delta_{\boldsymbol{\xi}} U_k + a \frac{\partial}{\partial \xi_1} U_k &= R_k, \qquad \boldsymbol{\xi} \neq \mathbf{0}, \\ U_k \to 0 \quad &\text{as } |\boldsymbol{\xi}| \to \infty, \end{aligned} \tag{3.7}$$

[4] To be consistent, the original variable $\mathbf{x}$ may be referred to as the *inner variable*. The transformation from inner variable to outer variable is a *contracting* rather than a *stretching* transformation.

where

$$R_1 = 0, \quad R_k = a \sum_{\nu=1}^{k-1} U_\nu \frac{\partial U_{k-\nu}}{\partial \xi_1}, \quad k \geqq 2.$$

Remark. The region $\bar{\Omega}_{\boldsymbol{\xi}}^C = E_2 \backslash \Omega_{\boldsymbol{\xi}}$ depends on ε and shrinks to a point as ε tends to zero. Thus, it might be expected that U_k satisfies the equation in (3.7) everywhere except at $\boldsymbol{\xi} = \mathbf{0}$. If one considers $\boldsymbol{\xi}$ as the original variable, then the boundary layer is the neighborhood of the origin (indeed, the neighborhood of the boundary $\partial\Omega_{\boldsymbol{\xi}}$).

Solutions of (3.7) are clearly not unique. Hence in order to account for the nonuniqueness, we need some kind of matching condition in the neighborhood of $\boldsymbol{\xi} = \mathbf{0}$. This can best be described from the construction of the term U_1.

For $k = 1$, the general solution of (3.7) is of the form[5]

$$U_1(\boldsymbol{\xi}) = -a_1 e^{-(a/2)\xi_1} K_0\left(\frac{a}{2}|\boldsymbol{\xi}|\right) \tag{3.8}$$

for arbitrary constant a_1. From (2.2) we have

$$U_1(\boldsymbol{\xi}) = a_1 \log|\boldsymbol{\xi}| + b_1 + O(|\boldsymbol{\xi}| \log|\boldsymbol{\xi}|) \quad \text{as } \boldsymbol{\xi} \to \mathbf{0}, \tag{3.9}$$

where

$$b_1 = a_1(\Gamma_0 - \log a).$$

To determine b_1, we first substitute (3.9) into the outer expansion (3.6) and obtain asymptotically,

$$U(\boldsymbol{\xi}; \varepsilon) = -a + \{a_1 \log|\boldsymbol{\xi}| + b_1 + O(|\boldsymbol{\xi}| \log|\boldsymbol{\xi}|)\}(\log\varepsilon)^{-1} + O(\log\varepsilon)^{-2}) \quad \text{as } \boldsymbol{\xi} \to \mathbf{0}. \tag{3.10}$$

Next we write (3.1) in the form

$$u\left(\frac{\boldsymbol{\xi}}{\varepsilon}; \varepsilon\right) \sim u_0\left(\frac{\boldsymbol{\xi}}{\varepsilon}; f\right) + \sum_{k=1}^{\infty} \left\{u_{A_k}\left(\frac{\boldsymbol{\xi}}{\varepsilon}; 0\right)\right\}(\log\varepsilon)^{-k}. \tag{3.11}$$

Now let ε tend to zero, with $\boldsymbol{\xi}$ fixed. Then the argument $\mathbf{x} = \boldsymbol{\xi}/\varepsilon$ becomes large and we substitute into (3.11) the asymptotic expansion (3.5). This yields

$$u\left(\frac{\boldsymbol{\xi}}{\varepsilon}; \varepsilon\right) \sim (-m_f - A_1) + \{A_1 \log|\boldsymbol{\xi}| - A_2\}(\log\varepsilon)^{-1} + O(\log\varepsilon)^{-2}. \tag{3.12}$$

The *matching principle* requires that coefficients of like powers of (log ε) should agree in (3.10) and (3.12), provided one neglects terms which tend to zero as $|\boldsymbol{\xi}| \to 0^+$. This yields

$$-m_f - A_1 = -a, \quad A_1 = a_1, \quad \text{and} \quad -A_2 = b_1. \tag{3.13}$$

Thus, we obtain $u_{A_1}(\mathbf{x}; 0)$ and $U_1(\boldsymbol{\xi})$. Now it is not difficult to see how the general

[5] The general solution contains also multiples of terms such as $e^{-(a/2)\xi_1} K_n((a/2)|\boldsymbol{\xi}|)$, where the K_n are modified Bessel functions of the second kind. Since $K_n((a/2)|\boldsymbol{\xi}|) = O(|\boldsymbol{\xi}|^{-n})$ as $\boldsymbol{\xi} \to \mathbf{0}$, these terms will be automatically rejected. This will be made clear when the matching principle is introduced.

matching principle can be formulated, provided one knows the behavior of the general solutions of (3.7) as $\boldsymbol{\xi}$ tends to zero.

The general solution of (3.7) admits the representation

$$U_k(\boldsymbol{\xi}) = -a_k\, e^{-(a/2)\xi_1} K_0\left(\frac{a}{2}|\boldsymbol{\xi}|\right) + D_k(\boldsymbol{\xi}), \qquad k \geqq 1, \tag{3.14}$$

where $D_1(\boldsymbol{\xi}) = 0$ and

$$D_k(\boldsymbol{\xi}) = a \sum_{n=1}^{k-1} \iint_{E_2} S(\boldsymbol{\xi}-\boldsymbol{\eta};\, 1) U_n(\boldsymbol{\eta}) \frac{\partial}{\partial \eta_1} U_{k-n}(\boldsymbol{\eta})\, d\boldsymbol{\eta}, \qquad k \geqq 2.$$

Here $S(\boldsymbol{\xi}-\boldsymbol{\eta};\, 1)$ is the fundamental singularity of $\tilde{\mathscr{L}}U = 0$ (see (2.1)); the a_k's are constants to be determined by the matching principle. In order to see that D_k indeed decays to zero at infinity, we need some estimates of D_k; furthermore, we also need some information about the singular behavior of D_k in the neighborhood of $\boldsymbol{\xi} = \mathbf{0}$.

To this end, we introduce the function $h_1 = h_1(|\boldsymbol{\xi}|)$ defined by

$$h_1(|\boldsymbol{\xi}|) = \begin{cases} \dfrac{1}{|\boldsymbol{\xi}|}, & 0 < |\boldsymbol{\xi}| \leqq 1, \\[2ex] \dfrac{1}{|\boldsymbol{\xi}|^{3/2}}, & |\boldsymbol{\xi}| > 1. \end{cases} \tag{3.15}$$

Remark. It is easy to verify that there exists a constant l such that

$$\left| \frac{\partial}{\partial \xi_1}\left(e^{-(a/2)\xi_1} K_0\left(\frac{a}{2}|\boldsymbol{\xi}|\right)\right) \right| \leqq l h_1\left(\frac{a}{2}|\boldsymbol{\xi}|\right).$$

LEMMA 3.1. *For* $\boldsymbol{\xi} \in E_2$, let

$$J(\boldsymbol{\xi}) = e^{-(a/2)\xi_1} \iint_{E_2} K_0\left(\frac{a}{2}|\boldsymbol{\xi}-\boldsymbol{\eta}|\right) K_0\left(\frac{a}{2}|\boldsymbol{\eta}|\right) h_1\left(\frac{a}{2}|\boldsymbol{\eta}|\right) d\boldsymbol{\eta}. \tag{3.16$_1$}$$

Then, J is continuous at $\boldsymbol{\xi} = \mathbf{0}$. Moreover, there exist constants H_1 and H_2 such that

$$|J(\boldsymbol{\xi})| \leqq \frac{H_1}{a^2} K_0\left(\frac{a}{2}|\boldsymbol{\xi}|\right) e^{-(a/2)\xi_1} \tag{3.16$_2$}$$

and

$$\left| \frac{\partial}{\partial \xi_1} J(\boldsymbol{\xi}) \right| \leqq \frac{H_2}{a} h_1\left(\frac{a}{2}|\boldsymbol{\xi}|\right). \tag{3.16$_3$}$$

The proof of the estimates $(3.16)_2$ and $(3.16)_3$ is tedious but straightforward and will, therefore, be omitted here. To establish the continuity, one can assume

$(a/2)|\boldsymbol{\xi}|<1$. Then consider the integral $(3.16)_1$ over the region $(a/2)|\boldsymbol{\eta}|\leqq 1$. By setting $f=K_0h_1$, we see that f is integrable for $(a/2)|\boldsymbol{\eta}|\leqq 1$. Hence the result follows immediately from the continuity of the integral

$$\iint_{(a/2)|\boldsymbol{\eta}|\leqq 1} \log\left(\frac{a}{2}|\boldsymbol{\xi}-\boldsymbol{\eta}|\right) f\left(\frac{a}{2}|\boldsymbol{\eta}|\right) d\boldsymbol{\eta}.$$

With the help of Lemma 3.1 and the remark below (3.15) it can be shown that $D_2(\boldsymbol{\xi})$ is continuous at $\boldsymbol{\xi}=\mathbf{0}$, and tends to zero as $|\boldsymbol{\xi}|\to\infty$. Hence by induction, it follows that $D_k(\boldsymbol{\xi})$ is continuous at $\boldsymbol{\xi}=\mathbf{0}$ and tends to zero as $|\boldsymbol{\xi}|\to\infty$. Thus, the solutions U_k of (3.7) satisfy

$$U_k(\boldsymbol{\xi})=a_k \log|\boldsymbol{\xi}|+b_k+O(|\boldsymbol{\xi}|\log|\boldsymbol{\xi}|) \quad \text{as } \boldsymbol{\xi}\to\mathbf{0} \tag{3.17}$$

where

$$b_k=a_k(\Gamma_0-\log a)+D_k(0).$$

We are now in a position to formulate the matching principle for the higher order terms.

MATCHING PRINCIPLE. *Determine the constants* A_k *of* (3.11) *and* a_k *of* (3.17) *so that the coefficients of* $\log|\boldsymbol{\xi}|$ *and the constant terms for corresponding powers of* $(\log\varepsilon)^{-1}$ *are equal.*

Comment. It was shown in [8] that for the model $(\hat{\mathrm{P}}_\varepsilon)$, the matching principle presented here may be considered as a simplified version of what is called the *asymptotic matching principle* in [6]. The same conclusion holds in the present case.

The matching principle gives a procedure for constructing formal inner and outer expansions. To establish the validity of this formal procedure is, in general, a difficult task, For the model problem of the ordinary differential equation $(\hat{\mathrm{P}}_\varepsilon)$, we have shown that the validity of the procedure can be completely verified. In the present case, because of technical difficulty, only partial justification is obtained. In § 6, we shall show that the process produces the correct first two terms in the expansions.

4. A priori estimates. We now consider the inhomogeneous problem

$$\begin{aligned} \mathscr{L}w &= \alpha(\phi\psi)_{x_1} && \text{in } \Omega, \\ w &= 0 && \text{on } \partial\Omega, \\ w &\to 0 && \text{as } |\mathbf{x}|\to\infty, \end{aligned} \tag{4.1}$$

where $\alpha=\varepsilon a$ and ϕ and ψ are functions of $\mathbf{x}$ satisfying certain conditions which will be specified later. We intend to majorize the solution of (4.1) in terms of something like the solution of the linear problem (P_0') in § 2. These estimates will be needed to establish the existence of a solution to the nonlinear problem (P_ε).

In view of the asymptotic behavior of the solutions of (P_0'), we introduce two auxiliary functions, $h_0 = h_0(r)$ and $\bar{h}_0 = \bar{h}_0(r)$, defined by:

$$h_0(r) = \begin{cases} \log \dfrac{2}{r}, & 0 < r \leqq 1, \\ \dfrac{1}{\sqrt{r}}, & r > 1, \end{cases} \tag{4.2}$$
$$\bar{h}_0(r) = \begin{cases} \log 2, & 0 < r \leqq 1, \\ h_0(r), & r > 1. \end{cases}$$

One can easily see that there exist constants l_1 and l_2 such that

$$l_1\, e^{-r} h_0(r) \leqq K_0(r) \leqq l_2\, e^{-r} h_0(r). \tag{4.3}$$

We now state the results as follows:

THEOREM 4. *Let $\phi(\mathbf{x}; \alpha)$ and $\psi(\mathbf{x}; \alpha)$ be functions of class C^2 in Ω and of class C^1 in $\bar{\Omega}$ such that, for $\mathbf{x} \in \bar{\Omega}$,*

$$|\psi(\mathbf{x}; \alpha)| \leqq B_0 h_0\left(\frac{\alpha}{2}|\mathbf{x}|\right) \quad \textit{and} \quad |\psi(\mathbf{x}; \alpha)| \leqq B_1 h_0\left(\frac{\alpha}{2}|\mathbf{x}|\right) \tag{4.4}$$

where B_0 and B_1 are constants. Then there exist an $\alpha_0 \in (0, 1)$, a constant $H = H(\Omega)$, and a solution $w(\mathbf{x}; \alpha)$ of (4.1) *such that*

$$|w(\mathbf{x}; \alpha)| \leqq H B_0 B_1 \bar{h}_0\left(\frac{\alpha}{2}|\mathbf{x}|\right), \qquad \mathbf{x} \in \bar{\Omega}, \tag{4.5}$$

for all $0 < \alpha \leqq \alpha_0$.

Remark. The proof of this theorem would be greatly facilitated by a knowledge of bounds for the derivatives of the Green's function. We were not able to obtain sufficiently sharp bounds and thus were forced into the rather complicated procedure of this section. However, for later use, a bound for the derivative of the Green's function will be given by Lemma 4.2 at the end of this section.

Our first observation is that

$$w^p(\mathbf{x}; \alpha) = \alpha \iint\limits_{\Omega} S(\mathbf{x} - \mathbf{y}; \alpha)(\phi\psi)_{y_1}\, d\mathbf{y}$$

is a particular solution of the equation in (4.1) provided that $(\phi\psi)_{y_1}$ is suitably restricted at infinity. An integration by parts yields

$$w^p(\mathbf{x}; \alpha) = \alpha \int_{\partial\Omega} S(\mathbf{x} - \mathbf{y}; \alpha) \cos(\mathbf{n}; y_1) \phi\psi\, dS_{\mathbf{y}} - \alpha \iint_{\Omega} \frac{\partial S}{\partial y_1}(\mathbf{x} - \mathbf{y}; \alpha) \phi\psi\, d\mathbf{y}. \tag{4.6}$$

Assume that (4.6) does indeed give a solution of the equation and let $w^h(\mathbf{x}; \alpha)$ be the solution of the problem (P_0') with $g(\mathbf{x})$ replaced by $-w^p$; the latter can be obtained as in § 2. Then the function $w = w^p + w^h$ is a solution of (4.1). To establish the inequality (4.5) it is necessary to estimate the solution of the problem (P_0'). We need a kind of maximum principle:

LEMMA 4.1. *Suppose* $0<\alpha<1$. *Then there exists a constant M depending on the geometry of* Ω *such that the solution* $v(\mathbf{x}; g; \alpha)$ *of* (P_0') *satisfies the inequality*:

$$|v(\mathbf{x}; g; \alpha)| \leqq M \|g\|_1 \frac{h_0(\alpha/2|\mathbf{x}|)}{|\log \alpha|} \tag{4.7}$$

uniformly on $\bar{\Omega}$, *where*

$$\|g\|_m = \max_{\mathbf{x}\in\partial\Omega} \left|\frac{d^k g}{dS^k}\right|, \qquad k \leqq m;$$

S is the arc-length.

The proof of Lemma 4.1 follows easily from the construction of $v(\mathbf{x}; g; \alpha)$. We omit the details. It is clear from (4.7) that w^h satisfies (4.5). Thus to complete the proof of Theorem 4, we need only show that w^p satisfies (4.5). The proof is technical and will be deferred to the Appendix. We remark, however, that the analysis is complicated by the existence of a region where the fundamental solution $S(\mathbf{x}-\mathbf{y}; \alpha)$ in (2.1) decays slowly. This is analogous to the Navier–Stokes equation [19] and represents a kind of *wake region* phenomenon.

To conclude the section, we state a lemma on an estimate of derivatives of the Green's function with respect to the domain Ω. This lemma will be needed for the existence proof.

LEMMA 4.2. *Let G be the Green's function of* $\mathscr{L}v = 0$ *for domain* Ω. *Then there exists a geometrical constant C such that for all* $\mathbf{x}, \mathbf{y} \in \Omega$,

$$\left|\frac{\partial G}{\partial y_1}\right| \leqq \left|\frac{\partial S}{\partial y_1}\right| + C\bar{h}_0\left(\frac{\alpha}{2}|\mathbf{x}|\right) \sup_{\mathbf{z}\in\partial\Omega} \left|\frac{\partial}{\partial y_1} S(\mathbf{z}-\mathbf{y}; \alpha)\right|. \tag{4.8}$$

Remark. The Green's function G can be written in the form

$$G(\mathbf{x}-\mathbf{y}; \alpha) = S(\mathbf{x}-\mathbf{y}; \alpha) + e^{-(\alpha/2)(x_1-y_1)} h(\mathbf{x}-\mathbf{y}; \alpha)$$

while an explicit form for the regular function h is given in [14]. The result (4.8) then follows easily with some manipulations. The details will be omitted here.

5. Existence theorem. In this section, we prove that there exists a solution to the problem (P_ε) for ε sufficiently small. Our approach is a variation of a technique due to Finn and Smith [5]. The idea here is to seek a solution of (P_ε) as the value at infinity, $-a$, plus a small perturbation. This leads to the consideration of the perturbed system for which the value at infinity tends to zero. Then it reduces the effect of the nonlinearity in the original equation and facilitates the construction of a solution. To be more specific, we consider the family of problems

$$\mathscr{L}v = \alpha\tau v v_{x_1} \quad \text{in } \Omega, \qquad \alpha = \varepsilon a,$$
$$(5.1) \qquad v = 1 + \frac{f}{a} \quad \text{on } \partial\Omega,$$
$$v \to 0 \qquad \text{as } |\mathbf{x}| \to \infty.$$

Here τ is a small parameter, $0 \leqq \tau \leqq 1$. Note that if $v(\mathbf{x}; \alpha; \tau)$ is a solution of (5.1), then

$$(5.2) \qquad u(\mathbf{x}; \varepsilon) = -a + av(\mathbf{x}; \alpha; 1)$$

is a solution of (P_ε). Hence, our object is to prove the existence of a solution of (5.1). At this point, it should be remarked that other methods can also be used for the existence proof but it should be constructive in order to study the asymptotic behavior of the solution in detail. We now state the results as follows:

THEOREM 5. *For α sufficiently small, there exists a solution $v(\mathbf{x}; \alpha; \tau)$ of (5.1) for $\tau \in [0, 1]$. This solution can be represented by an absolutely and uniformly convergent series*

$$(5.3)_1 \qquad v(\mathbf{x}; \alpha; \tau) = \sum_{n=0}^{\infty} v_n(\mathbf{x}; \alpha)\tau^n$$

where each $v_n(\mathbf{x}; \alpha)$ satisfies

$$\mathscr{L}v_n = \psi_n \quad \text{in } \Omega,$$
$$(5.3)_2 \qquad v_0 = 1 + \frac{f}{a}, \quad v_n = 0, \quad n \geqq 1, \quad \text{on } \partial\Omega,$$
$$v_n \to 0 \quad \text{as } |\mathbf{x}| \to \infty \text{ for all } n \geqq 0$$

with

$$\psi_0 = 0, \qquad \psi_n = \alpha \sum_{k=0}^{n-1} v_k \frac{\partial}{\partial x_1} v_{n-1-k}, \qquad n \geqq 1.$$

COROLLARY 5.1. *For ε sufficiently small, there exists a solution $u(\mathbf{x}; \varepsilon)$ of (P_ε), which can be represented as*

$$(5.4) \qquad u(\mathbf{x}; \varepsilon) = -a + a \sum_{n=0}^{\infty} v_n(\mathbf{x}; \varepsilon a).$$

The first step in the proof is to construct the series (5.3). v_0 can be constructed by the method in § 2, and hence by Lemma 4.1 we have

$$(5.5) \qquad |v_0| \leqq C_0 h_0\left(\frac{\alpha}{2}|\mathbf{x}|\right) \quad \text{for } \mathbf{x} \in \bar{\Omega},$$

where $C_0 = M_0/|\log \alpha|$, M_0 a constant. We note that

$$\sum_{k=0}^{n-1} v_k \frac{\partial v_{n-1-k}}{\partial x_1} = \frac{1}{2} \frac{\partial}{\partial x_1} \sum_{k=0}^{n-1} v_k v_{n-1-k}.$$

Thus if we use Theorem 4 and the fact that $\bar{h}_0((\alpha/2)|\mathbf{x}|) < h_0((\alpha/2)|\mathbf{x}|)$, we readily verify, by induction, that all the v_n's exist.

Our next step is to establish the convergence of the series (5.3) under the hypothesis of the theorem. This is based on (5.5) and the following lemma, which is an immediate consequence of Theorem 4.

LEMMA 5.1. *Let the sequence $\{C_n\}$ of constants be defined by*

$$C_{n+1} = H \sum_{k=0}^{n} C_k C_{n-k}, \qquad C_0 = \frac{M_0}{|\log \alpha|}. \tag{5.6}$$

Then

$$|v_n(\mathbf{x};\alpha)| \leqq C_n \bar{h}_0\left(\frac{\alpha}{2}|\mathbf{x}|\right), \qquad n \geqq 1,$$

uniformly on $\bar{\Omega}$.

It follows from Lemma 5.1 that the series

$$C(\tau;\alpha) = (\|v_0\| - C_0) + \sum_{k=0}^{\infty} C_k \tau^k \tag{5.7}$$

will dominate (5.3), since $\bar{h}_0((\alpha/2)|\mathbf{x}|)$ is bounded and less than one. Here $\|v_0\| = \sup_{x\in\Omega}|v_0(\mathbf{x};\alpha)|$, which is bounded by the maximum principle. By an analysis analogous to that in [8], we can show that

$$C_n \leqq \frac{\gamma_{n+1}(2HC_0)^n}{(n+1)!} C_0, \qquad n \geqq 0, \tag{5.8}$$

where

$$\gamma_{n+1} = \begin{cases} 1, & n = 0, \\ \prod_{k=2}^{n+1}(2k-3), & n \geqq 1. \end{cases}$$

Moreover the series (5.7) converges for $\tau < (4HC_0)^{-1}$. From the definition of C_0, this latter holds for all $\tau \in [0, 1]$, provided ε is sufficiently small. Consequently, (5.3) converges uniformly and absolutely.

Our final step is to demonstrate that the sum (5.3), $v(\mathbf{x};\alpha;\tau)$, satisfies the differential equation in (5.1). This can be facilitated by considering the corresponding integral equation. Our main task here is to establish the result:

LEMMA 5.2. *The function v defined by* (5.3) *is a solution of the integral equation*

$$v(\mathbf{x};\alpha;\tau) = v_0(\mathbf{x};\alpha) - \frac{\alpha}{2}\iint_{\Omega} v^2 \frac{\partial}{\partial y_1} G(\mathbf{x}-\mathbf{y};\alpha)\, d\mathbf{y}, \tag{5.9}$$

for the Green's function G.

Using (5.9), we find that it is straightforward to verify that $v(\mathbf{x};\alpha;\tau)$ has the required differentiability and that it satisfies (5.1).

To complete the proof of Theorem 5, we establish the result (5.9). First, observe that the v_n's can be written in the form

$$v_n(\mathbf{x};\alpha)=-\frac{\alpha}{2}\iint_\Omega \frac{\partial G}{\partial y_1}\left\{\sum_{k=0}^{n-1} v_k v_{n-1-k}\right\} d\mathbf{y}, \qquad n\geqq 1. \tag{5.10}$$

Hence, if we write

$$v^m=\sum_{n=0}^{m} v_n\tau^n, \tag{5.11}$$

then (5.10) implies that

$$v^m=v_0-\frac{\alpha\tau}{2}\iint_\Omega \frac{\partial G}{\partial y_1}\left\{\sum_{n=0}^{m-1}\sum_{k=0}^{n}(v_k v_{n-k})\tau^n\right\} d\mathbf{y}, \qquad m\geqq 1, \tag{5.12}$$

and

$$v=v_0-\frac{\tau\alpha}{2}\iint_\Omega v^2\frac{\partial G}{\partial y_1}d\mathbf{y}+(v-v^m)+\frac{\tau\alpha}{2}\iint_\Omega \frac{\partial G}{\partial y_1}\left\{v^2-\sum_{n=0}^{m-1}\left(\sum_{k=0}^{n} v_k v_{n-k}\right)\tau^n\right\} d\mathbf{y}.$$

Thus, to verify (5.9), it suffices to show that the last two terms on the right hand side can be made arbitrarily small for m sufficiently large. This is clear for the first term, $v-v^m$, from the uniform convergence of the series. However for the second term, we need a more involved procedure. This will proceed as follows.

Let $\Phi_m=v^2-\sum_{n=0}^{m-1}(\sum_{k=0}^{n} v_k v_{n-k})\tau^n$. For any fixed $\mathbf{x}\in\Omega$, we write

$$\alpha\iint_\Omega \left|\frac{\partial G}{\partial y_1}\right| |\Phi_m|\, d\mathbf{y}=\alpha\left\{\iint_{D_\rho(x)}+\iint_{\Omega\setminus D_\rho(x)}\right\}\left|\frac{\partial G}{\partial y_1}\right| |\Phi_m|\, d\mathbf{y}, \tag{5.13}$$

where $D_\rho(\mathbf{x})$ is a disc with $\mathbf{x}$ as center and radius ρ. If we choose ρ small enough, then for $\mathbf{y}\in D_\rho(\mathbf{x})$, $|\partial G/\partial y_1|$ will be dominated by some constant times $1/\alpha|\mathbf{x}-\mathbf{y}|$. Hence for any given $\mu>0$, there exists a ρ so small that

$$\alpha\iint_{D_\rho(x)}\left|\frac{\partial G}{\partial y_1}\right| |\Phi_m|\, d\mathbf{y}<\frac{\mu}{3}. \tag{5.14}$$

Now for this fixed ρ, by Lemma 4.2, it is possible to choose a number R so large that for all $|\mathbf{y}|\geqq R$ and $|\mathbf{x}-\mathbf{y}|\geqq\rho$,

$$\left|\frac{\partial G}{\partial y_1}\right|\leqq\left|\frac{\partial S(\mathbf{x}-\mathbf{y};\alpha)}{\partial y_1}\right|+C\left|\frac{\partial S(-\mathbf{y};\alpha)}{\partial y_1}\right| \tag{5.15}$$

where C is a constant depending only on the geometry of Ω. Then with ρ and R

fixed, we have

$$\alpha \iint_{\Omega \backslash D_\rho(x)} \left|\frac{\partial G}{\partial y_1}\right| |\Phi_m| \, d\mathbf{y} \leqq \alpha \left\{ \iint_{\substack{\rho \leqq |\mathbf{y}-\mathbf{x}| \\ |\mathbf{y}| \leqq R}} + \iint_{\substack{\rho \leqq |\mathbf{y}-\mathbf{x}| \\ |\mathbf{y}| > R}} \right\} \left|\frac{\partial G}{\partial y_1}\right| |\Phi_m| \, d\mathbf{y}.$$

In the region where $\rho \leqq |\mathbf{y}-\mathbf{x}|$ and $|\mathbf{y}| \leqq R$, since $\partial G/\partial y_1$ is regular, we obtain

$$\alpha \iint_{\substack{\rho \leqq |\mathbf{y}-\mathbf{x}| \\ |\mathbf{y}| \leqq R}} \left|\frac{\partial G}{\partial y_1}\right| dy \leqq M(\rho; R) \tag{5.16}$$

for some constant M. Then given any $\mu > 0$, there exists a number $N_1(\mu)$ such that for all $m \geq N_1(\mu)$, we have

$$|\Phi_m(\mathbf{y})| < \frac{\mu}{3M(\rho; R)}. \tag{5.17}$$

This is so because of the uniform convergence of the series $\sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} v_k v_{n-k}\right) \tau^n$. Therefore, we have from (5.16) and (5.17)

$$\alpha \iint_{\substack{\rho \leqq |\mathbf{y}-\mathbf{x}| \\ |\mathbf{y}| \leqq R}} \left|\frac{\partial G}{\partial y_1}\right| |\Phi_m| \, d\mathbf{y} < \frac{\mu}{3}. \tag{5.18}$$

In the region where $\rho \leqq |\mathbf{y}-\mathbf{x}|$ and $|\mathbf{y}| > R$, we obtain by (5.15) and (5.6)

$$\begin{aligned} \alpha \iint_{\substack{\rho \leqq |\mathbf{y}-\mathbf{x}| \\ |\mathbf{y}| > R}} \left|\frac{\partial G}{\partial y_1}\right| |\Phi_m| \, d\mathbf{y} \leqq \alpha \Bigg[\iint_{\substack{\rho \leqq |\mathbf{y}-\mathbf{x}| \\ |\mathbf{y}| > R}} \Bigg\{ \left| \frac{\partial S(\mathbf{x}-\mathbf{y}; \alpha)}{\partial y_1} \right| \\ + C \left| \frac{\partial S(-\mathbf{y}; \alpha)}{\partial y_1} \right| \Bigg\} \left\{ e^{-(\alpha/2) y_1} h_0\left(\frac{\alpha}{2}|\mathbf{y}|\right) \right\}^2 d\mathbf{y} \Bigg] \\ \cdot \sum_{n=m}^{\infty} \left(\sum_{k=0}^{n} C_k C_{n-k} \right) \tau^n. \end{aligned} \tag{5.19}$$

Now both terms in the square brackets are bounded independently of $\mathbf{x}$, say by M_1 (see (A.1)). Moreover, the series $\sum_{n=0}^{\infty} (\sum_{k=0}^{n} C_k C_{n-k}) \tau^n$ converges for $\tau \in [0, 1]$, for sufficiently small α. Hence given $\mu > 0$ there exists a number $N_2(\mu)$ such that for all $m \geqq N_2$,

$$\sum_{n=m}^{\infty} \left(\sum_{k=0}^{n} C_k C_{n-k} \right) \tau^n < \frac{\mu}{6M_1}. \tag{5.20}$$

Consequently, it follows from (5.14), (5.18), (5.19), and (5.20) that for any given

$\mu > 0$, there exists an $N = \max\{N_1, N_2\}$ such that

$$\left|\alpha \iint_\Omega \frac{\partial G}{\partial y_1}\left\{v^2 - \sum_{n=0}^{m-1}\sum_{k=0}^{n} v_k v_{n-k}\right\} d\mathbf{y}\right| < \mu$$

for all $m \geqq N$. This completes the proof of Lemma 5.2.

6. Asymptotic properties. From the constructed solution $u(\mathbf{x}; \varepsilon)$, one can now study the asymptotic behavior of the solution as ε tends to zero. We shall concentrate on the leading term v_0 of the series (5.3). In particular, we attempt to make some justification of the formal inner and outer expansions constructed in § 3. We begin with the following basic result.

LEMMA 6.1. *For ε sufficiently small and for any fixed integer $m \geqq 0$, the solution $u(\mathbf{x}; \varepsilon)$ defined by (5.4) satisfies the inequality:*

$$\left|u(\mathbf{x}; \varepsilon) - \left\{-a + a \sum_{j=0}^{m} v_j(\mathbf{x}; \varepsilon a)\right\}\right| \leqq \frac{d_{m+1}}{|(\log \varepsilon)^{m+2}|}\bar{h}_0\left(\frac{\varepsilon a}{2}|\mathbf{x}|\right) \tag{6.1}$$

uniformly on $\bar{\Omega}$, where $d_{m+1} < \infty$ is a constant independent of ε and $\mathbf{x}$.

The proof follows easily from Lemma 5.1. We omit the details.

Lemma 6.1 yields a kind of asymptotic development for the exact solution $u(\mathbf{x}; \varepsilon)$ of (P_ε). We see that since $\bar{h}_0((\varepsilon a/2)|\mathbf{x}|)$ is bounded and less than one, we have

$$u(\mathbf{x}; \varepsilon) = -a + a \sum_{j=0}^{m} v_j(\mathbf{x}; \varepsilon a) + O(\log \varepsilon)^{-(m+2)} \quad \text{as } \varepsilon \to 0^+ \tag{6.2}$$

uniformly on $\bar{\Omega}$. In particular, for $m = 0$ we have shown that

$$a v_0(\mathbf{x}; \varepsilon a) = -e^{-(\varepsilon a/2)x_1} \int_{\partial\Omega} \phi(\mathbf{y}) K_0\left(\frac{\varepsilon a}{2}|\mathbf{x} - \mathbf{y}|\right) dS_{\mathbf{y}},$$

where ϕ can be determined by the method in § 2. Then from Theorem 3, we arrive at two expansions. More precisely we have proved the following results:

THEOREM 6. *Let $\mathcal{D}$ be any compact subset of Ω and let $\mathcal{D}_\delta$ denote the region $\{\mathbf{x} \in E_2 : |\mathbf{x}| \geqq \delta\}$ for any $\delta > d/\varepsilon$ with $d = \sup\{|\mathbf{x}| : \mathbf{x} \in \partial\Omega\}$. Then we have,*

$$u(\mathbf{x}; \varepsilon) = u_0(\mathbf{x}; f) + u_{a-m_f}(\mathbf{x}; 0)(\log \varepsilon)^{-1} + O(\log \varepsilon)^{-2} \quad \textit{as } \varepsilon \to 0^+ \tag{6.3}$$

uniformly on $\mathcal{D}$; and

$$u(\mathbf{x}; \varepsilon) = -a - (a - m_f)\, e^{-(\varepsilon/2)ax_1} K_0\left(\frac{\varepsilon a}{2}|\mathbf{x}|\right)(\log \varepsilon)^{-1} + O(\log \varepsilon)^{-2} \quad \textit{as } \varepsilon \to 0^+ \tag{6.4}$$

uniformly on $\mathcal{D}_\delta$.

COROLLARY 6.1. *The solution $u(\mathbf{x}; \varepsilon)$ has the representation for $\mathbf{x} \in \bar{\Omega}$,*

$$u(\mathbf{x}; \varepsilon) = w(\mathbf{x}; \varepsilon) + v_\varepsilon(\varepsilon \mathbf{x}; \varepsilon) + z(\mathbf{x}; \varepsilon),$$

where

$$w(\mathbf{x};\varepsilon)=(1-\psi_\rho(\varepsilon\mathbf{x}))\{u_0(\mathbf{x};f)+u_{a-m_f}(\mathbf{x};0)(\log\varepsilon)^{-1}\},$$

$$v_\varepsilon(\varepsilon\mathbf{x};\varepsilon)=\psi_\rho(\varepsilon\mathbf{x})\left\{-a-(a-m_f)\,e^{-(\varepsilon/2)ax_1}K_0\left(\frac{\varepsilon a}{2}|\mathbf{x}|\right)(\log\varepsilon)^{-1}\right\},$$

$$z(\mathbf{x};\varepsilon)=O(\log\varepsilon)^{-2}\quad \text{as } \varepsilon\to 0^+$$

uniformly on $\bar{\Omega}$.

The proof of Theorem 2 then follows immediately from (6.3)–(6.4).

Appendix. Proof of Theorem 4. From Lemma 4.1, it is easy to see that the solution, w^h, of the corresponding homogeneous problem will satisfy the inequality

$$|w^h(\mathbf{x};\alpha)|\leqq M'\|w^p\|_1\bar{h}_0\left(\frac{\alpha}{2}|\mathbf{x}|\right)$$

in $\bar{\Omega}$ for some constant M'. Hence it suffices to consider the particular solution w^p in (4.6). To this end, we need the following two lemmas.

LEMMA A.1. *Suppose* $0<\alpha<1$. *Then there exists a constant* H_1, *depending only on* Ω, *such that*

$$\text{(A.1)}\qquad \begin{aligned}|J_1(\mathbf{x};\alpha)|&\equiv\left|\alpha\int_{\partial\Omega}S(\mathbf{x}-\mathbf{y};\alpha)\left\{h_0\left(\frac{\alpha}{2}|\mathbf{y}|\right\}^2 dS_{\mathbf{y}}\right.\\ &\leqq H_1\bar{h}_0\left(\frac{\alpha}{2}|\mathbf{x}|\right),\quad \text{for } \mathbf{x}\in\bar{\Omega}.\end{aligned}$$

Proof. Without loss of generality we may assume that $(\alpha/2)|\mathbf{y}|\leqq 1$ for $\mathbf{y}\in\partial\Omega$. Then the function $J_1(\mathbf{x};\alpha)$ is dominated by

$$\text{(A.2)}\qquad \bar{J}_1(\mathbf{x};\alpha)\equiv\left\{\alpha(\log\alpha)^2\int_{\partial\Omega}e^{-(\alpha/2)(x_1-y_1)}K_0\left(\frac{\alpha}{2}|\mathbf{x}-\mathbf{y}|\right)dS_{\mathbf{y}}\right\}.$$

For $(\alpha/2)|\mathbf{x}|>1$, clearly $\bar{J}_1(x;\alpha)$ can be made less than some constant times $\bar{h}_0((\alpha/2)|\mathbf{x}|)$, since $\mathbf{y}\in\partial\Omega$. For $(\alpha/2)|\mathbf{x}|\leqq 1$ the integral in (A.2) will be dominated by $|\log\alpha|$, and since $\alpha(\log\alpha)^3$ is bounded, the result still follows.

LEMMA A.2. *Suppose* $0<\alpha<1$. *Then there exists a constant* H_2, *depending only on* Ω, *such that*

$$\text{(A.3)}\qquad \begin{aligned}|J_2(\mathbf{x};\alpha)|&\equiv\left|\alpha\iint_{\Omega}\frac{\partial}{\partial y_1}S(\mathbf{x}-\mathbf{y};\alpha)\left\{h_0\left(\frac{\alpha}{2}|\mathbf{y}|\right)\right\}^2 d\mathbf{y}\right|\\ &\leqq H_2\bar{h}_0\left(\frac{\alpha}{2}|\mathbf{x}|\right),\quad \text{for } \mathbf{x}\in\bar{\Omega}.\end{aligned}$$

Before starting our proof, we record a lemma which we will need in the following. A proof of this lemma can be found in [6, p. 198].

LEMMA. *Let the numbers ρ and σ be such that $\rho<2$, $\sigma<2$, $\rho+\sigma>2$. Then there exists a constant C_1 depending on ρ and σ such that*

$$\iint_{E_2} \frac{d\mathbf{z}}{|\mathbf{x}-\mathbf{z}|^{\rho}|\mathbf{z}-\mathbf{y}|^{\sigma}} \leqq \frac{C_1}{|\mathbf{x}-\mathbf{y}|^{\rho+\sigma-2}} \tag{A.4}$$

for all $\mathbf{x}, \mathbf{y} \in E_2$ *and* $\mathbf{x} \neq \mathbf{y}$.

Throughout the proof, we denote constants by A_k. These constants may depend on the domain Ω and their values may change.

Proof of Lemma A.2. We split the domain Ω into subregions Ω' and Ω'' where $\Omega' = \{\mathbf{y}: (\alpha/2)|\mathbf{y}| \geqq 1\}$ and $\Omega'' = \{\mathbf{y}: (\alpha/2)|\mathbf{y}| < 1\}$. The corresponding parts of J_2 in (A.3) are then denoted respectively by J_2' and J_2''. To facilitate the proof, we first observe that for $|\boldsymbol{\eta}| \geqq 1$,

$$\left|\frac{\partial}{\partial \eta_1}\left\{e^{\eta_1} K_0(|\boldsymbol{\eta}|)\right\}\right| \leqq \text{const.}\ \frac{\exp\left(-|\boldsymbol{\eta}|(1-(\eta_1/|\boldsymbol{\eta}|))\right)}{|\boldsymbol{\eta}|^{1/2}}\left\{1-\frac{\eta_1}{|\boldsymbol{\eta}|}\right\},$$

which implies that $(\partial/\partial \eta_1)\{e^{\eta_1} K_0(|\boldsymbol{\eta}|)\}$ vanishes exponentially with $\boldsymbol{\eta}$ except in the region (the so-called "wake region" in fluid mechanics):

$$|\boldsymbol{\eta}|\left(1-\frac{\eta_1}{|\boldsymbol{\eta}|}\right) = O(1).$$

However, in this region, we have

$$1-\frac{\eta_1}{|\boldsymbol{\eta}|} = O\left(\frac{1}{|\boldsymbol{\eta}|}\right).$$

Hence we can conclude that for $|\boldsymbol{\eta}| \geqq 1$,

$$\left|\frac{\partial}{\partial \eta_1}\{e^{\eta_1} K_0(|\boldsymbol{\eta}|)\}\right| \leqq \frac{\text{const.}}{|\boldsymbol{\eta}|^{3/2}} \tag{A.5}$$

whether it is in the wake region or not.

Now let us consider J_2'. By changing variables,

$$|J_2'(\mathbf{x};\alpha)| = \frac{1}{2\pi}\left|\iint_{|\boldsymbol{\eta}|\geqq 1} \frac{\partial}{\partial \eta_1}\{e^{-(\xi_1-\eta_1)} K_0(|\boldsymbol{\xi}-\boldsymbol{\eta}|)\}\frac{1}{|\boldsymbol{\eta}|}\, d\boldsymbol{\eta}\right| \tag{A.6}$$

where $|\boldsymbol{\xi}| = (\alpha/2)|\mathbf{x}|$. We now consider two cases according to the location of $\boldsymbol{\xi}$:

Case 1. $0<|\boldsymbol{\xi}|\leqq 1$. Equation (A.6) implies that

$$\begin{aligned} |J_2'(\mathbf{x};\alpha)| &\leqq A_1\left\{\iint_{1\leqq|\boldsymbol{\eta}|\leqq 2}\left|\frac{\partial K_0(|\boldsymbol{\xi}-\boldsymbol{\eta}|)}{\partial \eta_1}\right| d\boldsymbol{\eta} + \iint_{2\leqq|\boldsymbol{\eta}|}\frac{d\boldsymbol{\eta}}{|\boldsymbol{\eta}|\,|\boldsymbol{\xi}-\boldsymbol{\eta}|^{3/2}}\right\} \\ &\leqq A_2 \leqq A_3 \bar{h}_0\left(\frac{\alpha}{2}|\mathbf{x}|\right). \end{aligned}$$

Case 2. $|\boldsymbol{\xi}| \geqq 1$. We can write, from (A.6),

$$|J_2'(\mathbf{x};\alpha)| \leqq \left| \left\{ \iint\limits_{\substack{|\boldsymbol{\eta}|\geqq 1 \\ |\boldsymbol{\xi}-\boldsymbol{\eta}|\leqq 1}} + \iint\limits_{\substack{|\boldsymbol{\eta}|\geqq 1 \\ |\boldsymbol{\xi}-\boldsymbol{\eta}|>1}} \right\} \frac{\partial}{\partial \eta_1}\{e^{-(\xi_1-\eta_1)}K_0(|\boldsymbol{\xi}-\boldsymbol{\eta}|)\}\frac{1}{|\boldsymbol{\eta}|}\,d\boldsymbol{\eta} \right|$$

$$\equiv |J_{21}'(\mathbf{x};\alpha)+J_{22}'(\mathbf{x};\alpha)|.$$

Here both $J_{21}'(\mathbf{x};\alpha)$ and $J_{22}'(\mathbf{x};\alpha)$ are dominated by the integral

$$\iint\limits_{|\boldsymbol{\eta}|\geqq 1} \frac{1}{|\boldsymbol{\xi}-\boldsymbol{\eta}|^{3/2}|\boldsymbol{\eta}|}\,d\boldsymbol{\eta}.$$

This follows easily from (A.5) and the fact that

$$\left|\frac{\partial}{\partial \eta_1}K_0(|\boldsymbol{\xi}-\boldsymbol{\eta}|)\right| \leqq |\boldsymbol{\xi}-\boldsymbol{\eta}|^{-3/2} \quad \text{for } |\boldsymbol{\xi}-\boldsymbol{\eta}|\leqq 1.$$

Hence an application of (A.4) yields the desired result that

$$|J_2'(\mathbf{x};\alpha)| \leqq A_1|\boldsymbol{\xi}|^{-1/2} = A_1\bar{h}_0\left(\frac{\alpha}{2}|\mathbf{x}|\right).$$

Next we consider J_2''. From the definition of h_0, we see that

$$\text{(A.7)} \qquad |J_2''(\mathbf{x};\alpha)| \leqq \iint\limits_{|\boldsymbol{\eta}|\leqq 1} \left|\frac{\partial}{\partial \eta_1}\{e^{-(\xi_1-\eta_1)}K_0(|\boldsymbol{\xi}-\boldsymbol{\eta}|)\}\right| \left(\log\frac{2}{|\boldsymbol{\eta}|}\right)^2 d\boldsymbol{\eta}$$

where $|\boldsymbol{\xi}| = (\alpha/2)|\mathbf{x}|$. Similarly, we consider two cases according to the location of $\boldsymbol{\xi}$:

Case 1. $|\boldsymbol{\xi}| \leqq 2$. Condition (A.7) implies that

$$|J_2''(\mathbf{x};\alpha)| \leqq A_1\left\{ \iint\limits_{|\boldsymbol{\eta}|\leqq 1} \frac{d\boldsymbol{\eta}}{|\boldsymbol{\xi}-\boldsymbol{\eta}|} + \iint\limits_{|\boldsymbol{\eta}|\leqq 1} \frac{1}{|\boldsymbol{\xi}-\boldsymbol{\eta}|}(\log|\boldsymbol{\eta}|)^2\,d\boldsymbol{\eta}\right\}.$$

This is clearly bounded, and we have

$$|J_2''(\mathbf{x};\alpha)| \leqq A_2\bar{h}_0\left(\frac{\alpha}{2}|\mathbf{x}|\right),$$

since for $(\alpha/2)|\mathbf{x}| \leqq 2$, $\bar{h}_0((\alpha/2)|\mathbf{x}|)$ is bounded by a nonzero constant.

Case 2. $|\boldsymbol{\xi}| > 2$. Condition (A.7) implies that

$$|J_2''(\mathbf{x};\alpha)| \leqq A_1 \iint\limits_{|\boldsymbol{\eta}|\leqq 1} \left|\frac{\partial}{\partial \eta_1}\{e^{-(\xi_1-\eta_1)}K_0(|\boldsymbol{\xi}-\boldsymbol{\eta}|)\}\right| (1+\log|\boldsymbol{\eta}|)^2\,d\boldsymbol{\eta}$$

$$\leqq A_2\frac{1}{\sqrt{|\boldsymbol{\xi}|}} \iint\limits_{|\boldsymbol{\eta}|<1} \frac{1}{\sqrt{1-|\boldsymbol{\eta}|/|\boldsymbol{\xi}|}}(1+\log|\boldsymbol{\eta}|)^2\,d\boldsymbol{\eta}$$

$$\leqq A_3\bar{h}_0\left(\frac{\alpha}{2}|\mathbf{x}|\right).$$

Now if we collect all the estimates, we find that the proof of Lemma A.2 is complete.

Lemmas A.1 and A.2 and condition (4.6) yield the estimate

$$|w^p(\mathbf{x};\alpha)| \leqq B_0 B_1 (H_1 + H_2)\bar{h}_0\left(\frac{\alpha}{2}|\mathbf{x}|\right), \qquad \mathbf{x} \in \bar{\Omega}.$$

By a similar argument one can show that there exists a constant H_3 such that $\|w^p\|_1 \leqq B_0 B_1 H_3$. Consequently, the result (4.5) follows if we let $H = H_1 + H_2 + MH_3$. This completes the proof of Theorem 4.

Acknowledgments. The author wishes to express his appreciation to Professor Wolfgang Wendland for his encouragment and hospitality during the author's stay in Germany. Thanks are also due to Professor Richard C. MacCamy for his advice throughout the development of the work.

REFERENCES

[1] F. P. Bretherton, *Slow viscous motion round a cylinder in a simple shear*, J. Fluid Mech., 12 (1962), pp. 591–613.

[2] J. D. Cole, *Perturbation Methods in Applied Mathematics*, Blaisdell, Waltham, MA, 1968.

[3] W. Eckhaus, *Matched Asymptotic Expansions and Singular Perturbation*, North-Holland, Amsterdam, 1973.

[4] R. Finn and D. R. Smith, *On the linearized hydrodynamical equations in two dimensions*, Arch. Rational Mech. Anal., 25 (1967), pp. 1–23.

[5] ———, *On the stationary solutions of the Navier–Stokes equations in two dimensions*, Ibid., 25 (1967), pp. 26–39.

[6] L. E. Fraenkel, *On the method of matched asymptotic expansions. Part I: A matching principle; Part II: Some applications of the composite series; Part III: Two boundary-value problems*, Proc. Cambridge Philos. Soc., 65 (1969), pp. 209–284.

[7] G. Hellwig, *Partial Differential Equations*, Blaisdell, New York, 1964.

[8] G. C. Hsiao, *Singular perturbations for a nonlinear differential equation with a small parameter*, this Journal, 4 (1973), pp. 283–301.

[9] G. C. Hsiao and R. C. MacCamy, *Solution of boundary value problems by integral equations of the first kind*, SIAM Rev., 15 (1973), pp. 687–705.

[10] R. P. Kanwal, *Applications of the technique of matched asymptotic expansions to various fields of mathematical physics*, Proc. Symp. Analyt. Methods in Math. Phys., Indiana University Press, Bloomington, Indiana, 1969, pp. 479–485.

[11] S. Kaplan, *Low Reynolds number flow past a circular cylinder*, J. Math. Mech., 6 (1957), pp. 595–603.

[12] P. A. Lagerstrom, *Some recent developments in the theory of singular perturbations*, Problems in Analysis: A Symposium in Honor of S. Bochner, Princeton University Press, Princeton, NJ, 1970, pp. 261–271.

[13] N. Levinson, *The first boundary value problem for* $\varepsilon\Delta u + Au_x + Bu_y + Cu = D$ *for small* ε, Ann. of Math., 51 (1950), pp. 428–445.

[14] R. C. MacCamy, *Low frequency acoustic oscillations*, Quart. Appl. Math., 23 (1965), pp. 247–255.

[15] N. I. Mushelishvili, *Singular Integral Equations*, P. Noordhoff, Groningen, The Netherlands, 1953.

[16] R. E. O'Malley, Jr., *Topics in singular perturbations*, Advances in Mathematics, vol. 2, Academic Press, New York, 1968, pp. 365–470.

[17] J. R. A. Pearson and I. Proudman, *Expansions at small Reynolds numbers for the flow past a sphere and a circular cylinder*, J. Fluid Mech., 2 (1957), pp. 237–262.

[18] S. ROSENBLAT AND J. SHEPHERD, *On the asymptotic solutions of the Lagerstrom model equation*, SIAM J. Appl. Math., 29 (1975), pp. 110–120.

[19] D. R. SMITH, *Estimates at infinity for stationary solutions of the Navier–Stokes equations in two dimensions*, Arch. Rational Mech. Anal., 20 (1965), pp. 341–372.

[20] M. VAN DYKE, *Perturbation Methods in Fluid-Dynamics*, Academic Press, New York, 1964.

[21] M. I. VISIK AND L. A. LYUSTERNIK, *Regular degeneration and boundary layer for linear differential equations with small parameter*, Uspehi Mat. Nauk, 12 (1957), no. 5 (77), pp. 3–122; Amer. Math. Soc. Transl., Ser. 2, 20 (1962), pp. 239–364.

[22] W. WASOW, *Asymptotic solution of boundary value problems for the differential equation* $\Delta u + \lambda(\partial u/\partial x) = \lambda f(x, y)$, Duke Math. J., 11 (1944), pp. 405–415.

[23] ———, *Asymptotic Expansions for Ordinary Differential Equations*, Interscience, New York, 1965.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

EXISTENCE AND ASYMPTOTIC STABILITY OF SOLUTIONS OF AN ABSTRACT INTEGRODIFFERENTIAL EQUATION WITH APPLICATIONS TO VISCOELASTICITY*

SARP ADALI†

Abstract. By using certain energy estimates, the existence of a unique solution for an abstract integrodifferential equation in function space is established. Then, the asymptotic behavior of solutions of the equation, which represents the abstract form of the dynamical equations of viscoelasticity theory, is investigated under conditions which are mechanically realistic in the framework of the theory of viscoelasticity. The recently developed theory of compact processes seems the most appropriate in this respect. The use of the invariance principle for the compact process generated by the equation under investigation, combined with the existence of a Lyapunov functional, leads to the proof that the solutions tend to zero as time goes to infinity. In the last part of the paper, the results are applied to the equations of viscoelasticity and the mechanical interpretation of the assumptions is given.

1. Introduction. In this paper we study the problems of existence, uniqueness and asymptotic stability for a class of abstract integrodifferential equations in function space. The results are then applied to viscoelasticity equations.

Let H_0, H_1 and H_2 be real Hilbert spaces with norms $\|\cdot\|_0$, $\|\cdot\|_1$ and $\|\cdot\|_2$, respectively, and such that $H_2 \subset H_1 \subset H_0$ algebraically and topologically. We define another space H_{-1} as the dual of H_1 via the inner product $\langle\cdot,\cdot\rangle$ of H_0. Thus, H_{-1} will be the completion of H_0 under the norm

$$\|W\|_{-1} \equiv \sup_{V \in H_1} \frac{|\langle W, V\rangle|}{\|V\|_1} \tag{1.1}$$

and $\langle\cdot,\cdot\rangle$ is extended onto $H_{-1} \times H_1$ as a continuous bilinear form. We will assume further that the injection of H_i into H_{i-1}, $i = 0, 1, 2$, is compact.

We consider the following history value problem:

$$\frac{d}{dt}(\rho(t)\dot{u}(t)) + C(t)u(t) + \int_{-\infty}^{t} G(t-\tau, t)u(\tau)\,d\tau = f(t) \tag{1.2}$$

for $t \in [s, \infty)$, where s is a given parameter, with the history

$$u(\tau+s) = v(\tau), \qquad \tau \in (-\infty, 0].^{1} \tag{1.3}$$

Here $\rho(t)$ is a self-adjoint operator in H_0 for every $t \in (-\infty, +\infty)$; $C(t)$, for fixed t, and $G(\xi, t)$, for fixed ξ and t, are bounded linear operators from H_1 to H_{-1}. We note that equation (1.2) represents the abstract form of the dynamical equations of viscoelasticity theory with (1.3) specifying the history and the independent variable t denoting the time. The viscoelastic body considered here has the property that $C(t)$, $G(\xi, t)$ and $\rho(t)$ approach time-independent limiting

* Received by the editors May 24, 1973, and in final revised form September 16, 1976.

† Department of Mathematics, Middle East Technical University, Ankara, Turkey. Now at Council for Scientific and Industrial Research, National Research Institute for Mathematical Sciences, Pretoria, Republic of South Africa.

[1] $\dot{u}(t)$ denotes the first derivative of $u(t)$.

values at large times as in the cases of many polymers. In this paper we are interested in the existence and asymptotic behavior of solutions of (1.2), (1.3) under conditions which are realistic in the framework of the theory of viscoelasticity and preferably have definite mechanistic interpretations. The investigation will answer the question of whether a Boltzmann type dependence of the stress on the history of deformation, when this history is specified up to a certain time, will induce a damping mechanism with the solutions uniquely determined.

Previously, the invariance principle has been exploited successfully in the investigation of the asymptotic behavior of solutions of evolutionary equations which generate dynamical systems [11], [5] or compact processes [6]. A similar approach seems to be the most appropriate for the abovementioned problem from the viewpoint of mechanics. Similar problems in viscoelasticity have been considered by Dafermos [5] and by MacCamy [13] for the case where $C(t)$, $G(\xi, t)$ and the density were time-independent, and they proved that solutions decay to zero as time goes to infinity under suitable conditions. Dafermos [4] has also investigated a similar initial value problem in function space and established a set of sufficient conditions for the asymptotic stability of its solutions. The methods in [5] and [13] for establishing the asymptotic stability fail in our case for the reason that both depend on kernels of convolution type.

In this paper, we investigate the asymptotic behavior of solutions of the equation (1.2) in the framework of the theory of compact processes developed by Dafermos [3]. The observation that (1.2), (1.3) generate a compact process which is, in the terminology of [3], asymptotically a dynamical system and the use of invariance principle combined with the existence of a Lyapunov functional for this process enable us to prove the asymptotic stability of the solutions of (1.2), (1.3) under a set of sufficient conditions which is different from, and in some applications [5] weaker than the corresponding set in [4]. For example, the assumption of convexity of $G(\xi, t)$ in [4] which does not admit a mechanistic interpretation is dropped and so is the positive-definiteness of $\partial G/\partial t$. Nevertheless, the results differ in form rather than in essence from their counterparts in [4]. Another advantage of the new method is that it emphasizes the history rather than the initial value problem, and consequently it leads in a natural way to function spaces of fading memory type. This is important from the viewpoint of mechanics since these spaces constitute a natural setting of viscoelasticity theory [2].

We first prove the existence of a unique solution $u(t)$ for the history value problem (1.2), (1.3) in § 2 by using certain energy estimates. Edelstein and Gurtin [10] and Odeh and Tadjbakhsh [14] studied the uniqueness problem for the classical viscoelasticity equations similar in form to (1.2) assuming the positivity of a constant density and the definiteness of the elastic modulus $C(t)$. We give the proof of an existence theorem for equation (1.2) essentially under the same conditions; however our guidance here will be the work of Dafermos [4], where he discusses the existence and uniqueness questions for an abstract Volterra equation taken as an initial value problem. Indeed, in establishing our results we adapt an existence theorem from [4].

The compact process generated by (1.2), (1.3) is studied in § 3. After constructing a Lyapunov functional for this process, we give our main result about the asymptotic stability of zero solutions in § 4.

In the last section we apply our results to the equations of linear viscoelasticity;

$$
(1.4)\quad \begin{aligned}\frac{\partial}{\partial t}\left(\rho(\mathbf{x}, t)\frac{\partial u_i(\mathbf{x}, t)}{\partial t}\right) = \frac{\partial}{\partial x_j}\bigg(C_{ijkl}(\mathbf{x}, t)\frac{\partial u_k(\mathbf{x}, t)}{\partial x_l} \\ + \int_{-\infty}^{t} G_{ijkl}(\mathbf{x}, t-\tau, t)\frac{\partial u_k(\mathbf{x}, \tau)}{\partial x_l}\,d\tau\bigg) + f_i(\mathbf{x}, t).\end{aligned}
$$

[2]

The reduction of equation (1.4) to the form (1.2) will be postponed to § 5. In this form, equations (1.4) represent the dynamical equations for a viscoelastic material when small deformations are superposed on a large deformation history and/or when there is an aging process taking place in the material. The asymptotic behavior of the solutions of (1.4) is considered when the initial deformation path tends to a stationary state and/or the aging process tends to stop as time goes to infinity. In § 5, after applying the results of the previous sections to viscoelasticity equations (1.4), we give the mechanical interpretation of various assumptions made in the paper.

2. Existence and uniqueness of solutions. Throughout this work we will assume that

1. $$C(t), \dot{C}(t) \in L^\infty((-\infty, \infty); \mathscr{L}(H_1; H_{-1})),$$ [3]
$$C(t), \dot{C}(t) \in L^\infty((-\infty, \infty); \mathscr{L}(H_2; H_0)).$$

Moreover

$$(2.1)\qquad \langle C(t)w, v\rangle = \langle C(t)v, w\rangle \quad \text{for all } v, w \in H_1, \quad \text{and } t \in (-\infty, \infty),$$

$$(2.2)\qquad \langle C(t)w, w\rangle \geqq K\|w\|_1^2 \quad \text{for all } w \in H_1, \quad \text{and } t \in (-\infty, \infty).$$

2. $$G(\xi, t) \in C^0([0, \infty)\times(-\infty, \infty); \mathscr{L}(H_2; H_0)) \cap L^\infty([0, \infty) \times(-\infty, \infty); \mathscr{L}(H_2; H_0)),$$
$$G(\xi, t), G_\xi(\xi, t), G_t(\xi, t) \in C^0([0, \infty)\times(-\infty, \infty); \mathscr{L}(H_1; H_{-1})) \cap L^\infty([0, \infty)\times(-\infty, \infty); \mathscr{L}(H_1; H_{-1})).$$ [4]

Furthermore, for fixed ξ and t,

$$(2.3)\qquad \langle G(\xi, t)w, v\rangle = \langle G(\xi, t)v, w\rangle \quad \text{for all } v, w \in H_1,$$

$$(2.4)\qquad \langle G(\xi, t)w, w\rangle \leqq 0 \quad \text{for all } w \in H_1.$$

3. $$\rho(t), \dot{\rho}(t), \ddot{\rho}(t), \dddot{\rho}(t) \in C^0((-\infty, \infty); \mathscr{L}(H_0; H_0)) \cap L^\infty((-\infty, \infty); \mathscr{L}(H_0, H_0)),$$

$$(2.5)\qquad \langle \rho(t)w, w\rangle \geqq \rho_0\|w\|_0^2, \quad \rho_0 > 0, \quad \text{for all } w \in H_0, \quad t \in [0, \infty),$$

---

[2] $\mathbf{x}$ denotes a point in the three-dimensional Euclidean space E^3. The summation convention is employed throughout the paper and i, j, k, l take the values 1, 2, 3.

[3] $\mathscr{L}(H_1; H_{-1})$ stands for the space of bounded linear operators from H_1 to H_{-1}.

[4] The subscripts denote differentiation with respect to that independent variable.

(2.6) $\langle \dot{\rho}(t)w, w\rangle \geqq 0$ for all $w \in H_0$, $t \in [0, \infty)$,

(2.7) $\langle \ddot{\rho}(t)w, w\rangle \geqq 0$ for all $w \in H_0$, $t \in [0, \infty)$.

(2.8) $\langle \dddot{\rho}(t)w, w\rangle \geqq 0$ for all $w \in H_0$, $t \in (-\infty, \infty)$,

4. $f(t), \dot{f}(t) \in L^{\infty}((-\infty, \infty); H_0)$, $f(t), \dot{f}(t) \in L^1([0, \infty); H_0)$.

We first state a lemma from Dafermos [5].

LEMMA 2.1. *Let $f(t) \in L^1(0, \infty)$. Then there exists an increasing function $p(t) \in C^0[0, \infty)$ with $p(0)=1$, $p(t) \to \infty$ as $t \to \infty$, such that $f(t)p(t) \in L^1(0, \infty)$.*

It follows that we can assume the existence of a decreasing "influence function" $h(t) \in C^0[0, \infty)$ with the following properties:

$$h(0)=1, \qquad h(0) \to 0 \quad \text{as } t \to \infty, \tag{2.9}$$

$$\int_0^{\infty} [\|G(\xi, t)\|_{\mathscr{L}(H_1; H_{-1})} + \|G_{\xi}(\xi, t)\|_{\mathscr{L}(H_1; H_{-1})}]h^{-2}(\xi)\, d\xi < M < \infty \tag{2.10}$$

for every $t \in (-\infty, \infty)$. After fixing some influence function $h(t)$ with the above properties, we proceed to the definition of some Banach spaces:

DEFINITION 2.1. By $\mathscr{C}_k$, $k=0, 1$, we denote the Banach space of functions $w(\tau) \in C^k((-\infty, 0]; H_1) \cap C^{k+1}((-\infty, 0]; H_0)$ such that

$$|||w|||_{\mathscr{C}_k} \equiv \sum_{i=0}^{k} \sup_{(-\infty,0]} [h(-\tau)\|\overset{(i)}{w}(\tau)\|_1] + \sum_{i=0}^{k+1} \sup_{(-\infty,0]} [h(-\tau)\|\overset{(i)}{w}(\tau)\|_0] < \infty.\text{[5]} \tag{2.11}$$

It is clear that the above norm attaches greater weight to the recent than the distant past, in accordance with the ideas of fading memory.[6]

DEFINITION 2.2. By $\mathscr{B}_k$, $k=0, 1$, we denote the Banach space of functions

$$w(\tau) \in C^k((-\infty, 0]; H_2) \cap C^{k+1}((-\infty, 0]; H_1) \cap C^{k+2}((-\infty, 0]; H_0)$$

such that

$$|||w|||_{\mathscr{B}_k} \equiv \sum_{i=0}^{k} \sup_{(-\infty,0]} \|\overset{(i)}{w}(\tau)\|_2 + \sum_{i=0}^{k+1} \sup_{(-\infty,0]} \|\overset{(i)}{w}(\tau)\|_1 + \sum_{i=0}^{k+2} \sup_{(-\infty,0]} \|\overset{(i)}{w}(\tau)\|_0 < \infty. \tag{2.12}$$

With these definitions we have $\mathscr{B}_1 \subset \mathscr{B}_0 \subset \mathscr{C}_1 \subset \mathscr{C}_0$ algebraically and topologically. We also have the following lemma from Dafermos [5].

LEMMA 2.2. *For $k=0, 1$, the injection of $\mathscr{B}_k$ into $\mathscr{C}_k$ is compact.*

Next we state two theorems which will be used later in the proof of existence and uniqueness of solutions.

Consider the initial value problem

$$\frac{d}{dt}(\rho(t)\dot{u}(t)) = C(t)u(t) - \int_s^t G(t-\tau, t)u(\tau)\, d\tau + f(t) \tag{2.13}$$

---

[5] By $\overset{(i)}{w}(\tau)$ we denote the ith derivative of $w(\tau)$. For simplicity we will write $\dot{w}(\tau)$ for $\overset{(1)}{w}(\tau)$ and $\ddot{w}(\tau)$ for $\overset{(2)}{w}(\tau)$.

[6] An account of the theories of fading memory in viscoelasticity can be found in [2] and [16].

for $t \geqq s$ and with the initial values

$$u(s)=u_0 \in H_1, \qquad \dot u(s)=\dot u_0 \in H_0. \tag{2.14}$$

Here s is some fixed real number. Let $f(t)=f^{(1)}(t)+f^{(2)}(t)$ and

$$f^{(1)}(t) \in L^1([s,\infty); H_0), \qquad f^{(2)} \in L^1([s,\infty); H_{-1}). \tag{2.15}$$

THEOREM 2.1. *There exists a unique solution* $u(t)$ *of* (2.13), (2.14) *and*

$$u(t) \in C^0([s,\infty); H_1), \qquad \dot u(t) \in C^0([s,\infty); H_0). \tag{2.16}$$

Furthermore,

$$\|u(t)\|_1+\|\dot u(t)\|_0 \leqq \sqrt{2c}\, e^{ct/2}\Big\{\|u_0\|_1+\|\dot u_0\|_0 + \sqrt{\frac{c}{2}}\int_s^t [\|f^{(1)}(\tau)\|_0+\|f^{(2)}(\tau)\|_{-1}]\, d\tau\Big\}. \tag{2.17}$$

Proof. We form the $\langle\cdot,\cdot\rangle$ product of (2.13) with $\dot u(t)$ and integrate over (s,t) and observe the following relation:

$$\int_s^t \Big\langle \frac{d}{d\tau}(\rho(\tau)\dot u(\tau)), \dot u(\tau)\Big\rangle d\tau = \frac{1}{2}\langle \rho(t)\dot u(t), \dot u(t)\rangle - \frac{1}{2}\langle \rho(s)\dot u(s), \dot u(s)\rangle + \frac{1}{2}\int_s^t \langle \dot\rho(\tau)\dot u(\tau), \dot u(\tau)\rangle\, d\tau. \tag{2.18}$$

Recalling (2.6), we can complete the rest of the proof along the same lines as that of Theorem 2.3 in [4].

Now we impose additional smoothness assumptions on the initial conditions, namely,

$$u(s)=u_0 \in H_2, \qquad \dot u(s)=\dot u_0 \in H_1. \tag{2.19}$$

Let $\ddot u(s) \in H_0$ be defined by

$$\ddot u(s)=\rho^{-1}(s)[-\dot\rho(s)\dot u(s)+C(s)u_0+f(s)].$$

We also define

$$G^{(1)}(t-\tau,t) \equiv \int_\tau^t G_t(t-\xi,t)\, d\xi + G(0,t) - \dot C(t), \tag{2.20}$$

$$f_1(t) \equiv \dot f(t) - \int_s^t G_t(t-\xi,t)u(s)\, d\xi - G(0,t)u(s) + \dot C(t)u(t). \tag{2.21}$$

Let $f_1(t)=f_1^{(1)}(t)+f_1^{(2)}(t)$ and

$$f_1^{(1)}(t) \in L^1([s,\infty); H_0), \qquad f_1^{(2)}(t), \dot f_1^{(2)} \in L^1([s,\infty); H_{-1}). \tag{2.22}$$

With these definitions, differentiation of (2.13) gives

$$\frac{d}{dt}(\dot\rho(t)\dot u(t)+\rho(t)\ddot u(t)) = C(t)\dot u(t) - \int_s^t G^{(1)}(t-\tau,t)\dot u(\tau)\, d\tau + f_1(t). \tag{2.23}$$

Now we are ready to state

THEOREM 2.2. *There is a unique solution* $u(t)$ *of* (2.13), (2.19) *and*

$$\dot{u}(t) \in C^1([s, \infty); H_1), \qquad \ddot{u}(t) \in C^0([s, \infty); H_0), \tag{2.24}$$

$$\begin{aligned} \|\dot{u}(t)\|_1 + \|\ddot{u}(t)\|_0 \leqq \sqrt{2c}\, e^{ct/2} \Big\{ \|\dot{u}_0\|_1 + \|\ddot{u}_0\|_0 + \int_s^t [\|f_1^{(2)}(\tau)\|_{-1} + \|\dot{f}_1^{(2)}(\tau)\|_{-1}]\, d\tau \\ + \sqrt{\frac{c}{2}} \int_s^t [\|f_1^{(1)}(\tau)\|_0 + \|\dot{f}_1^{(2)}(\tau)\|_{-1}]\, d\tau \Big\}. \end{aligned} \tag{2.25}$$

Proof. We form the $\langle \cdot, \cdot \rangle$ product of (2.23) with $\ddot{u}(t)$ and integrate over (s, t). We observe the following relation:

$$\begin{aligned} &\int_s^t \left\langle \frac{d}{d\tau}(\dot{\rho}(\tau)\dot{u}(\tau) + \rho(\tau)\ddot{u}(\tau)),\, \ddot{u}(\tau) \right\rangle d\tau \\ &= \frac{1}{2}\langle \rho(t)\ddot{u}(t), \ddot{u}(t)\rangle + \frac{1}{2}\langle \ddot{\rho}(t)\dot{u}(t), \dot{u}(t)\rangle - \frac{1}{2}\langle \rho(s)\ddot{u}(s), \ddot{u}(s)\rangle \\ &\quad - \frac{1}{2}\langle \ddot{\rho}(s)\dot{u}(s), \dot{u}(s)\rangle + \frac{3}{2}\int_s^t \langle \dot{\rho}(\tau)\ddot{u}(\tau), \ddot{u}(\tau)\rangle\, d\tau - \frac{1}{2}\int_s^t \langle \dddot{\rho}(\tau)\dot{u}(\tau), \dot{u}(\tau)\rangle\, d\tau. \end{aligned} \tag{2.26}$$

Recalling (2.6), (2.7), (2.8) we observe that the rest of the proof follows the same lines as that of Theorem 2.3 in Dafermos [4].

Now we state and prove two theorems about the existence and uniqueness of the solutions of (1.2), (1.3).

THEOREM 2.3. *For* $v \in \mathscr{C}_k$, $k = 0, 1$, *and* $T > s$, *there exists a unique* $u(t)$ *such that* $u(t) \in C^k((-\infty, T]; H_1) \cap C^{k+1}((-\infty, T]; H_0)$ *which satisfies* (1.3) *on* $(-\infty, s]$ *and* (1.2) *on* $[s, T]$. *Furthermore,*

$$\sum_{i=0}^{k} \sup_{[s,T]} \|\overset{(i)}{u}(t)\|_1 + \sum_{i=0}^{k+1} \sup_{[s,T]} \|\overset{(i)}{u}(t)\|_0 \leqq C_1 |||v|||_{\mathscr{C}_k} + C_2 \sum_{i=0}^{k} \sup_{[s,\infty)} \|\overset{(i)}{f}(t)\|_0 \tag{2.27}$$

where C_1 *and* C_2 *are independent of* v.

Proof. Let us set

$$F(t) \equiv f(t) - \int_{-\infty}^{s} G(t-\tau, t) v(\tau - s)\, d\tau. \tag{2.28}$$

With this definition we can rewrite the equation (1.2) in the following form;

$$\frac{d}{dt}(\rho(t)\dot{u}(t)) + C(t)u(t) + \int_s^t G(t-\tau, t)u(\tau)\, d\tau = F(t). \tag{2.29}$$

With the initial data specified as

$$u(s) = v(0), \qquad \dot{u}(s) = \dot{v}(0) \tag{2.30}$$

we have the same initial value problem as considered in Theorem 2.1. We have the following obvious bounds for $u(s)$, $\dot{u}(s)$ and $\ddot{u}(s)$;

$$\|u(s)\|_1 = \|v(0)\|_1 \leqq \sup_{(-\infty,0]} [h(-\tau)\|v(\tau)\|_1], \tag{2.31}$$

$$\|\dot{u}(s)\|_1 = \|\dot{v}(0)\|_1 \leqq \sup_{(-\infty,0]} [h(-\tau)\|\dot{v}(\tau)\|_1], \tag{2.32}$$

$$\|\ddot{u}(s)\|_0 = \|\ddot{v}(0)\|_0 \leqq \sup_{(-\infty,0]} [h(-\tau)\|\ddot{v}(\tau)\|_0]. \tag{2.33}$$

Consider now the decomposition;

$$F(t) = F^{(1)}(t) + F^{(2)}(t), \tag{2.34}$$

$$F^{(1)}(t) \equiv f(t) \in L^\infty((-\infty, \infty); H_0), \tag{2.35}$$

$$F^{(2)}(t) \equiv -\int_{-\infty}^{s} G(t-\tau, t)v(\tau - s)\, d\tau. \tag{2.36}$$

We have the following bounds for $F^{(1)}(t)$ and $F^{(2)}(t)$:

$$\|F^{(1)}(t)\|_0 = \|f(t)\|_0 \leqq \sup_{[s,\infty)} \|f(t)\|_0, \tag{2.37}$$

$$\begin{aligned} \|F^{(2)}(t)\|_{-1} &= \left\| -\int_{-\infty}^{s} G(t-\tau, t)v(\tau - s)\, d\tau \right\|_{-1} \\ &\leqq \left(\int_{-\infty}^{s} \|G(t-\tau, t)\|_{\mathscr{L}(H_1; H_{-1})} h^{-1}(-\tau + s)\, d\tau \right) \\ &\quad \cdot \sup_{(-\infty,0)} [h(-\tau)\|v(\tau)\|_1]. \end{aligned} \tag{2.38}$$

If we set $t - \tau = \xi$ and note that $t - s \geqq 0$ and $h^{-1}(\xi) \geqq h^{-1}(\xi - t + s)$, after a simple computation we obtain

$$\begin{aligned} \sup_{[s,T]} \|F^{(2)}(t)\|_{-1} \leqq \int_0^\infty & \sup_{t \in (-\infty,\infty)} \|G(\xi, t)\|_{\mathscr{L}(H_1; H_{-1})} h^{-1}(\xi)\, d\xi \\ & \cdot \sup_{(-\infty,0]} [h^{-1}(-\tau)\|v(\tau)\|_1]. \end{aligned} \tag{2.39}$$

In the same way, we can obtain bounds on $\dot{F}^{(1)}(t)$ and $\dot{F}^{(2)}(t)$;

$$\|\dot{F}^{(1)}(t)\|_0 = \|\dot{f}(t)\|_0 \leqq \sup_{[s,\infty)} \|\dot{f}(t)\|_0, \tag{2.40}$$

$$\begin{aligned} \sup_{[s,T]} \|\dot{F}^{(2)}(t)\|_{-1} \leqq \Bigg(\int_0^\infty & \Bigg[\sup_{t \in (-\infty,\infty)} \|G_\xi(\xi, t)\|_{\mathscr{L}(H_1; H_{-1})} \\ & + \sup_{t \in (-\infty,\infty)} \|G_t(\xi, t)\|_{\mathscr{L}(H_1; H_{-1})} \Bigg] h^{-1}(\xi)\, d\xi \Bigg) \\ & \cdot \sup_{(-\infty,0]} [h^{-1}(-\tau)\|v(\tau)\|_1]. \end{aligned} \tag{2.41}$$

Now using (2.17) and (2.25) we obtain the estimate (2.27). Assertion of the theorem thus follows from Theorems 2.1 and 2.2.

THEOREM 2.4. *For $v \in \mathscr{B}_0$ and $T > s$, there exists a unique $u(t)$ such that*

$$u(t) \in C^0((-\infty, T]; H_2) \cap C^1((-\infty, T], H_1) \cap C^2((-\infty, T]; H_0)$$

which satisfies (1.3) *on* $(-\infty, s]$ *and* (1.2) *on* $[s, T]$. *Furthermore,*

$$\sup_{[s,T]} \|u(t)\|_2 + \sum_{i=0}^{1} \sup_{[s,T]} \|\overset{(i)}{u}(t)\|_1 + \sum_{i=0}^{2} \sup_{[s,T]} \|\overset{(i)}{u}(t)\|_0 \leqq C_1 |||v|||_{\mathscr{B}_0} + C_2 \sum_{i=0}^{1} \sup_{[s,\infty)} \|\overset{(i)}{f}(t)\|_0. \tag{2.42}$$

Proof. Since $\mathscr{B}_0 \subset \mathscr{C}_1$ we have $v \in \mathscr{C}_1$. Then we can deduce the following results from Theorem 2.3:

1. $u(t) \in C^1((-\infty, T]; H_1) \cap C^2((-\infty, T]; H_0)$,
2. $u(t)$ is the unique solution of (1.2) on $[s, T]$ and satisfies equation (1.3) on $(-\infty, s]$.

There remains to show $u(t) \in C^0([s, T]; H_2)$. First we establish the following estimate by using the inequality (2.27) for $k = 1$;

$$\begin{aligned} \left\| \frac{d}{dt}(\rho(t)\dot{u}(t)) \right\|_0 &\leqq \|\ddot{u}(t)\|_0 \|\rho(t)\|_0 + \|\dot{u}(t)\|_0 \|\dot{\rho}(t)\|_0 \\ &\leqq \sup_{[s,\infty)} (\|\rho(t)\|_0 + \|\dot{\rho}(t)\|_0) \left[C_1 |||v|||_{\mathscr{B}_0} + C_2 \sum_{i=0}^{1} \sup_{[s,\infty)} \|\overset{(i)}{f}(t)\|_0 \right]. \end{aligned} \tag{2.43}$$

Since this estimate is good for any $t \in [s, T]$, we can conclude

$$\begin{aligned} \sup_{[s,T]} \left\| \frac{d}{dt}(\rho(t)\dot{u}(t)) \right\|_0 &\leqq \sup_{[s,\infty)} (\|\rho(t)\|_0 + \|\dot{\rho}(t)\|_0) \\ &\quad \cdot \left[C_1 |||v|||_{\mathscr{B}_0} + C_2 \sum_{i=0}^{1} \sup_{[s,\infty)} \|\overset{(i)}{f}(t)\|_0 \right]. \end{aligned} \tag{2.44}$$

Now we obtain an estimate for $F(t)$, defined by (2.28), in H_0; that is, we visualize $G(\xi, t)$ as a linear operator from H_2 to H_0. Using the fact that

$$\left[\sup_{(-\infty,0]} \|v(\tau)\|_2 \right] \leqq |||v|||_{\mathscr{B}_0}, \tag{2.45}$$

we have

$$\sup_{[s,T]} \|F(t)\|_0 \leqq \sup_{[s,\infty)} \|f(t)\|_0 + \left(\int_0^\infty \sup_{t \in [s,\infty)} \|G(\xi, t)\|_{\mathscr{L}(H_2; H_0)} \, d\xi \right) |||v|||_{\mathscr{B}_0}. \tag{2.46}$$

To complete the proof we shall need the following lemma:

LEMMA 2.3. *For $g(t) \in C^0([s, T]; H_0)$ there exists a solution $w(t) \in C^0([s, t]; H_2)$ of the integral equation*

$$C(t)w(t) + \int_s^t G(t-\tau, t) w(\tau) \, d\tau = g(t) \tag{2.47}$$

on $[s, T]$. *Furthermore*

$$\sup_{[s,T]} \|w(t)\|_2 \leqq C \sup_{[s,T]} \|g(t)\|_0 \tag{2.48}$$

where C *is independent of* $g(t)$.

Proof. We set $w_0 \equiv 0$ and apply the standard Volterra iteration scheme to obtain the following expression for $w_n(t)$:

$$w_n(t) \equiv C^{-1}(t)g(t) - C^{-1}(t) \int_s^t G(t-\tau, t)w_{n-1}(\tau)\, d\tau, \tag{2.49}$$

for $n = 1, 2, \cdots$.

By taking the difference $w_{n+1}(t) - w_n(t)$ and noting the assumption (2.2), after a simple computation we obtain the following estimate:

$$\|w_{n+1}(\xi) - w_n(\xi)\|_2 \leqq \frac{M\xi}{K} \sup_{[s,\xi]} \|w_n(\tau) - w_{n-1}(\tau)\|_2 \tag{2.50}$$

where $\xi \in [s, T]$ and

$$M_\xi \equiv \int_s^\xi \|G(\xi - \tau, \xi)\|_{\mathscr{L}(H_2;\, H_0)}\, d\tau. \tag{2.51}$$

Since $M_\xi = \sup_{[s,\xi]} M_\xi$, we have

$$\sup_{[s,\xi]} \|w_{n+1}(t) - w_n(t)\|_2 \leqq \frac{M\xi}{K} \sup_{[s,\xi]} \|w_n(\tau) - w_{n-1}(\tau)\|_2. \tag{2.52}$$

From (2.49) and (2.2) we deduce

$$\sup_{[s,\xi]} \|w_1(t)\|_2 \leqq \frac{1}{K} \sup_{[s,\xi]} \|g(t)\|_0. \tag{2.53}$$

From this result and (2.25) it follows that

$$\sup_{[s,\xi]} \|w_{n+1}(t) - w_n(t)\|_2 \leqq \frac{M_\xi^n}{K^n} \sup_{[s,\xi]} \|g(t)\|_0. \tag{2.54}$$

If we choose ξ sufficiently small so that $M_\xi/K < 1$, $\{w_n(t)\}$ becomes a Cauchy sequence in $C^0([s, \xi]; H_2)$, and thus converges to some $w(t) \in C^0([s, \xi]; H_2)$ as $n \to \infty$ and $w(t)$ is a solution of the integral equation (2.47) on $[s, \xi]$. We also have the following estimate for $w(t)$:

$$\sup_{[s,\xi]} \|w(t)\|_2 \leqq \sum_{n=0}^{\infty} \sup_{[s,\xi]} \|w_{n+1}(t) - w_n(t)\|_2 \leqq \frac{1}{1 - M\xi/K} \sup_{[s,\xi]} \|F(t)\|_0. \tag{2.55}$$

To obtain the estimate (2.48), we extend $w(t)$ onto $[s, T]$ by a step by step argument.

Proof of Theorem 2.4 (continued). We have already established by (2.44) and (2.46) that $\|(d/dt)(\rho(t)\dot{u}(t))\|_0$ and $\|F(t)\|_0$ are uniformly bounded, and their sum gives $g(t)$. This implies that $|u(t)\|_2$ is also uniformly bounded due to the estimate (2.48) of Lemma 2.3 and this establishes (2.42). The assertion of the theorem then follows from Lemma 2.3 and Theorem 2.3.

Remark 2.1. Suppose that

$$g(t) \in C^0([s, \infty); H_0) \cap L^\infty([s, \infty); H_0)$$

and define $M_\infty(\xi)$ in the following way:

$$M_\infty(\xi) \equiv \int_0^\infty \|G(\zeta, \xi)\|_{\mathscr{L}(H_2; H_0)} \, d\zeta. \tag{2.56}$$

If $M_\infty(\xi)/K < 1$ for $\xi \in (-\infty, \infty)$, then the sequence $\{w_n(t)\}$ given by (2.49) will converge to $w(t)$ uniformly on $[s, \infty)$ and in place of (2.55) we have

$$\sup_{[s,\infty)} \|w(t)\|_2 \leqq \frac{1}{1 - \dfrac{M_\infty(\xi)}{K}} \sup_{[s,\infty)} \|g(t)\|_0. \tag{2.57}$$

3. The process generated by equation (1.2). In this section we make some additional assumptions;

$$\dot{\rho}(t) \in L^1([0, \infty); \mathscr{L}(H_0; H_0)), \tag{3.1}$$

$$\dot{C}(t) \in L^1([0, \infty); \mathscr{L}(H_1; H_{-1})). \tag{3.2}$$

These conditions are related to the assumed time-independent behavior of $\rho(t)$ and $C(t)$ at large times. Indeed, as consequences of (3.1) and (3.2) we have

$$\rho(t) \xrightarrow{H_0} \rho, \qquad t \to \infty, \tag{3.3}$$

$$C(t) \xrightarrow{\mathscr{L}(H_1; H_{-1})} C, \qquad t \to \infty. \tag{3.4}$$

For the relaxation function $G(\xi, t)$, we assume that for each fixed $t \in (-\infty, \infty)$

$$G(\cdot, t) \in L^1([0, \infty); \mathscr{L}(H_1; H_{-1})). \tag{3.5}$$

Also in agreement with the ideas of fading memory, we require that for each fixed $t \in (-\infty, \infty)$

$$G_\xi(\cdot, t), G_t(\cdot, t) \in L^1([0, \infty); \mathscr{L}(H_1; H_{-1})). \tag{3.6}$$

Furthermore, it is assumed that

$$\int_0^\infty \|G(\xi, t)\|_{\mathscr{L}(H_1; H_{-1})} \, d\xi \leqq k < \infty. \tag{3.7}$$

The last two conditions on $G(\xi, t)$ express the assumption that $G(\xi, t)$ approaches a steady state as $t \to \infty$. Indeed, as a consequence of (3.6) and (3.7) we have

$$\int_0^\infty \|G(\xi, t) - G(\xi)\|_{\mathscr{L}(H_1; H_{-1})} \, d\xi \to 0 \quad \text{as } t \to \infty, \tag{3.8}$$

where $G(\xi)$ is the limiting value of $G(\xi, t)$ as $t \to \infty$. It is clear that ρ, C and $G(\xi)$ satisfy (2.1), (2.2), (2.3), (2.4) and (2.5).

We now consider the mapping[7]

$$\omega: \mathbb{R} \times \mathscr{C}_0 \times \mathbb{R}^+ \to \mathscr{C}_0 \tag{3.9}$$

[7] We employ the notation $\mathbb{R} = (-\infty, \infty)$, $\mathbb{R}^+ \equiv [0, \infty)$.

which sends (s, v, ξ), $s \in \mathbb{R}$, $v \in \mathscr{C}_0$, $\xi \in \mathbb{R}^+$, onto $\omega^s(v, \xi) \in \mathscr{C}_0$ defined by

$$\omega^s(v, \xi)(\tau) \equiv u(s+\xi+\tau), \qquad \tau \in (-\infty, 0], \tag{3.10}$$

where $u(t)$ satisfies (1.2) on $[s, \xi]$ and (1.3) on $(-\infty, s]$. Theorem 2.3 implies that ω is well defined.

THEOREM 3.1. *The map* $\omega: \mathbb{R} \times \mathscr{C}_0 \times \mathbb{R}^+ \to \mathscr{C}$ *which sends* (s, v, ξ) *to* $\omega^s(v, \xi)$ *is a process on* $\mathscr{C}_0$ (see [3, Definition 2.1]).

Proof. (i) We have by (3.10) and (1.3)

$$\omega^s(v, 0)(\tau) = u(s+\tau) = v(\tau) \quad \text{for all } t \in \mathbb{R}, \tag{3.11}$$

i.e.,

$$\omega^s(v, 0) = v \quad \text{for all } v \in \mathscr{C}_0, \quad s \in \mathbb{R}. \tag{3.12}$$

(ii) From the definition of the map ω it follows that

$$\omega^s(v, \xi+\zeta)(\tau) = u(s+\xi+\zeta+\tau) = \omega^{s+\xi}(\omega^s(v, \xi), \zeta)(\tau) \tag{3.13}$$

for all $v \in \mathscr{C}_0$, $s \in \mathbb{R}$; $\xi, \zeta \in \mathbb{R}^+$.

(iii) In the proof of Theorem 2.3 we have obtained uniform bounds for $F^{(1)}(t)$ and $\dot{F}^{(1)}(t)$ in H_0 and for $F^{(2)}(t)$ and $\dot{F}^{(2)}(t)$ in H_{-1} such that the constants C_1 and C_2 in the estimate (2.27) are independent of s. Then using (2.27) we obtain

$$\begin{aligned} |||\omega^s(v, \xi)(\tau)|||_{\mathscr{C}_0} = &\sup_{\tau \in (-\infty, 0]} [h(-\tau)\|u(s+\xi+\tau)\|_1] \\ &+ \sum_{i=0}^{1} \sup_{\tau \in (-\infty, 0]} [h(-\tau)\|\overset{(i)}{u}(s+\xi+\tau)\|_0] \\ &\leqq C_1|||v|||_{\mathscr{C}_0} + C_2 \sup_{[s, \infty)} \|f(t)\|_0. \end{aligned} \tag{3.14}$$

Since (1.2) is linear, (3.14) implies that for any fixed $\xi \in \mathbb{R}^+$, the one-parameter family of maps $\omega^s(\cdot, \xi): \mathscr{C}_0 \to \mathscr{C}_0$, with parameter s, is equicontinuous. This completes the proof.

On account of Theorem 2.4 the restriction of ω to $\mathbb{R} \times \mathscr{B}_0 \times \mathbb{R}^+$ is also a map

$$\omega: \mathbb{R} \times \mathscr{B}_0 \times \mathbb{R}^+ \to \mathscr{B}_0. \tag{3.15}$$

THEOREM 3.2. *The map* $\omega: \mathbb{R} \times \mathscr{B}_0 \times \mathbb{R}^+ \to \mathscr{B}_0$ *which sends* (s, u, ξ) *to* $\omega^s(v, \xi)$ *with* $s \in \mathbb{R}$, $v \in \mathscr{B}_0$, $\xi \in \mathbb{R}^+$ *is a process on* $\mathscr{B}_0$.

Proof. It is the same as the proof of Theorem 3.1, except we use estimate (2.42) instead of (2.27) for establishing equicontinuity.

Consider now the map $\bar{\omega}: \mathscr{C}_0 \times \mathbb{R}^+ \to \mathscr{C}_0$ which sends (v, ξ), $v \in \mathscr{C}_0$, $\xi \in \mathbb{R}^+$, onto $\bar{\omega}(v, \xi) \in \mathscr{C}_0$ defined by

$$\bar{\omega}(v, \xi)(\tau) \equiv y(\xi+\tau), \qquad \tau \in (-\infty, 0], \tag{3.16}$$

where $y(t)$ satisfies

$$\frac{d}{dt}(\rho \dot{y}(t)) + Cy(t) + \int_{-\infty}^{t} G(t-\tau)y(\tau)\, d\tau = 0 \tag{3.17}$$

on $[0, \xi]$ and

$$y(\tau) = v(\tau) \quad \text{on } (-\infty, 0]. \tag{3.18}$$

Here ρ, C and $G(\xi)$ are defined by (3.3), (3.4) and (3.8), respectively; that is, they are the limits of $\rho(t)$, $C(t)$ and $G(\xi, t)$ as t goes to infinity. It can be shown that the map $\bar{\omega}$ is a dynamical system ([3, Definition 4.1]). For a proof we refer to [5, p. 304]. The connection between ω and $\bar{\omega}$ is established by the following theorem:

THEOREM 3.3. *The process ω is asymptotically a dynamical system with asymptotic hull* $\{\bar{\omega}\}$ (see [3, Definitions 2.7 and 4.4]).

Proof. We have to show that for any fixed $v \in \mathscr{C}_0$ and $\xi \in \mathbb{R}^+$,

$$\omega^s(v, \xi) \xrightarrow{\mathscr{C}_0} \bar{\omega}(v, \xi), \quad \text{as } s \to \infty. \tag{3.19}$$

Forming the difference $\omega^s(v, \xi) - \bar{\omega}(v, \xi)$ we observe that

$$\omega^s(v, \xi)(\tau) - \bar{\omega}(v, \xi)(\tau) = W(s + \xi + \tau) \tag{3.20}$$

with

$$W(s + \xi + \tau) = u(s + \xi + \tau) - y(\xi + \tau) \tag{3.21}$$

where $u(t)$ is the solution of (1.2), (1.3) and $y(t)$ is the solution of (3.17), (3.18). By subtracting (3.17) from (1.2) we deduce that $W(r)$ satisfies

$$\begin{aligned} &\frac{d}{dr}(\rho(r)\dot{W}(r)) + C(r)W(r) + \int_s^r G(r-\tau, r)W(\tau)\,d\tau \\ &\quad = f(r) - [\rho(r) - \rho]\ddot{y}(r-s) - \dot{\rho}(r)\dot{y}(r-s) - [C(r) - C]y(r-s) \\ &\quad\quad - \int_{-\infty}^r [G(r-\tau, r) - G(r-\tau)]y(\tau - s)\,d\tau \end{aligned} \tag{3.22}$$

on $[s, \infty)$ and

$$W(\tau) = 0 \quad \text{for } \tau \in (-\infty, s]. \tag{3.23}$$

We now define $f_0^{(1)}(r)$ and $f_0^{(2)}(r)$ by

$$f_0^{(1)}(r) \equiv f(r) - [\rho(r) - \rho]\ddot{y}(r-s) - \dot{\rho}(r)\dot{y}(r-s), \tag{3.24}$$

$$f_0^{(2)}(r) \equiv -[C(r) - C]y(r-s) - \int_{-\infty}^r [G(r-\tau, r) - G(r-\tau)]y(\tau - s)\,d\tau. \tag{3.25}$$

By using (2.17), (3.23) and the fact that $l\|u(t)\|_1 \geqq \|u(t)\|_0$ for all $t \in (-\infty, \infty)$ we deduce

$$\begin{aligned} &\|W(r)\|_0 + \|W(r)\|_1 + \|\dot{W}(r)\|_0 \\ &\quad \leqq \sqrt{2c}\, e^{c(r-s)/2} \Big\{ \sqrt{\frac{c}{2}} \int_s^r \|f(\eta)\|_0\,d\eta + \int_s^r \|[\rho(\eta) - \rho]\ddot{y}(\eta - s)\|_0\,d\eta \\ &\quad\quad + \int_s^r \|\dot{\rho}(\eta)\dot{y}(\eta - s)\|_0\,d\eta + \int_s^r \|[C(\eta) - C]y(\eta - s)\|_{-1}\,d\eta \\ &\quad\quad + \int_s^r \Big\| \int_{-\infty}^r [G(\eta - \tau_1, \eta) - G(\eta - \tau_1)]y(\eta - s)\,d\tau_1 \Big\|_{-1} d\eta \Big\}. \end{aligned} \tag{3.26}$$

By noting that (3.17), (3.18) is a special case of (1.2), (1.3) we apply Theorem 2.3 and obtain the estimate

$$\sup_{[0,\xi]} \|y(t)\|_1 + \sum_{i=0}^{1} \sup_{[0,\xi]} \|\overset{(i)}{y}(t)\|_0 \leqq C_1 |||v|||_{\mathscr{C}_0} \tag{3.27}$$

where ξ is any fixed positive number.

We now let $r = s + \xi + \tau$, $\eta = s + \sigma$ and $s - \tau_1 = -\tau_2$ and take the supremum of both sides of the inequality (3.26) over $\tau \in (-\infty, 0]$ to obtain

$$\sup_{\tau\in(-\infty,0]} (\|W(s+\xi+\tau)\|_0 + \|W(s+\xi+\tau)\|_1 + \|\dot W(s+\xi+\tau)\|_0)$$

$$\begin{aligned} &\leqq \sqrt{2c}\, e^{c\xi/2}\Big\{\sqrt{\frac{c}{2}}\int_0^\xi \|f(s+\sigma)\|_0\, d\sigma + \int_0^\xi \|\rho(s+\sigma)-\rho\|_0 \|\ddot y(\sigma)\|_0\, d\sigma \\ &\quad + \int_0^\xi \|\dot\rho(s+\sigma)\|_0 C_1 |||v|||_{\mathscr{C}_0}\, d\sigma + \int_0^\xi \|C(s+\sigma)-C\|_{\mathscr{L}(H_1;\, H_{-1})} C_1 |||v|||_{\mathscr{C}_0}\, d\sigma \\ &\quad + \int_0^\xi \int_{-\infty}^\xi \|G(\sigma-\tau_2, \sigma+s) - G(\sigma-\tau_2)\|_{\mathscr{L}(H_1;\, H_{-1})} C_1 |||v|||_{\mathscr{C}_0}\, d\tau_2\, d\sigma\Big\}. \end{aligned} \tag{3.28}$$

When we pass to the limit as $s \to \infty$, the right-hand side of this inequality goes to zero on account of the integrability of f, $\dot\rho$ and conditions (3.3), (3.4), (3.8) on $\rho(t)$, $C(t)$ and $G(\xi, t)$. Recalling (3.20), we see that this result implies that

$$\lim_{s\to\infty} |||\omega^s(v, \xi)(\tau) - \bar\omega(v, \xi)(\tau)|||_{\mathscr{C}_0} = 0, \tag{3.29}$$

that is, we obtain (3.19).

4. Asymptotic behavior of solutions. Throughout this section we will assume that

$$\langle G_t(\xi, t)w, w\rangle \geqq 0 \quad \text{for all } w \in H_1, \quad t \in (-\infty, \infty), \quad \xi \in [0, \infty), \tag{4.1}$$

$$\langle G_\xi(\xi, t)w, w\rangle \geqq 0 \quad \text{for all } w \in H_1, \quad t \in (-\infty, \infty), \quad \xi \in [0, \infty), \tag{4.2}$$

$$\frac{1}{K}\int_0^\infty \|G(\xi, t)\|_{\mathscr{L}(H_2;\, H_0)}\, d\xi < 1 \quad \text{for } t \in (-\infty, \infty), \tag{4.3}$$

$$\langle A(t)w, w\rangle \geqq a_0\|w\|_1^2 \quad \text{for } a_0 > 0, \quad \text{for all } w \in H_1, \quad t \in (-\infty, \infty), \tag{4.4}$$

$$\Big\langle\Big(\dot C(t) + \int_0^\infty G_t(\xi, t)\, d\xi\Big)w, w\Big\rangle \leqq 0 \quad \text{for all } w \in H_1, \quad t \in (-\infty, \infty), \tag{4.5}$$

where K is the same constant as in inequality (2.2) and A is defined by the following identity:

$$A(t) \equiv C(t) + \int_{-\infty}^t G(t-\tau, t)\, d\tau. \tag{4.6}$$

We note that (4.1) is weaker than the corresponding assumption (3.6) of [4]. Furthermore, convexity condition (3.8) on $G(\xi, t)$ of [4] is not present here. On the other hand, assumption (4.3) made here has no counterpart in [4]. But we remark that (4.3) and (4.4) are in the same spirit and in some applications these two conditions coincide as shown in § 5 and [5].

In this section we will prove the stability theorem for the solutions of (1.2), (1.3).

Let a functional $V^s(v)$ on $\mathscr{C}_0 \times R$ be defined by

$$V^s(v) \equiv \left[\frac{1}{2\rho_0} + E^s(v)\right] \exp\left(+\int_s^\infty \|f(\tau)\|_0 \, d\tau\right) \tag{4.7}$$

where $v \in \mathscr{C}_0$, $s \in \mathbb{R}$ and

$$\begin{aligned} E^s(v) \equiv & \frac{1}{2}\langle \rho(s)\dot{v}(0), \dot{v}(0)\rangle + \frac{1}{2}\langle A(s)v(0), v(0)\rangle \\ & -\frac{1}{2}\int_{-\infty}^0 \langle G(-\xi, s)(v(0)-v(\xi)), (v(0)-v(\xi))\rangle \, d\xi. \end{aligned} \tag{4.8}$$

The functional $V^s(v)$ was constructed for the classical Volterra equation by Levin [12]; he attributes the motivation to Volterra [15].

Due to (2.5), (4.4), (2.4) and the integrability of $f(t)$,

$$V^s(v) \geqq \frac{1}{2\rho_0} + \frac{1}{2}\rho_0\|\dot{v}(0)\|_0^2 + \frac{1}{2}a_0\|v(0)\|_1^2 \geqq 0. \tag{4.9}$$

Next we compute

$$\dot{V}^s(v) \equiv \limsup_{\varepsilon \to 0^+} \frac{1}{\varepsilon}(V^{s+\varepsilon}(\omega^s(v, \varepsilon)) - V^s(v)) \tag{4.10}$$

by temporarily assuming that $v \in \mathscr{C}_1$. We add and subtract the term $[1/(2\rho_0) + E^{s+\varepsilon}(\omega^s(v, \varepsilon))] \exp(+\int_s^\infty \|f(\tau)\|_0 \, d\tau)$ to the right-hand side of (4.10) and after grouping terms together we obtain

$$\begin{aligned} \dot{V}^s(v) = \limsup_{\varepsilon \to 0^+} \frac{1}{\varepsilon}\Big\{ & \left[\frac{1}{2\rho_0} + E^{s+\varepsilon}(\omega^s(v, \varepsilon))\right] \\ & \cdot \left[\exp\left(+\int_{s+\varepsilon}^\infty \|f(\tau)\|_0 \, d\tau\right) - \exp\left(+\int_s^\infty \|f(\tau)\|_0 \, d\tau\right)\right] \\ & + [E^{s+\varepsilon}(\omega^s(v, \varepsilon)) - E^s(v)] \exp\left(+\int_s^\infty \|f(\tau)\|_0 \, d\tau\right)\Big\}. \end{aligned} \tag{4.11}$$

On account of (2.5), (4.4) and (2.4), we have

$$E^s(v) \geqq \tfrac{1}{2}\rho_0\|\dot{v}(0)\|_0^2 + \tfrac{1}{2}a_0\|v(0)\|_1^2. \tag{4.12}$$

From (1.3), (2.6), (4.1), (4.2), (4.5), (4.12) and the inequality

$$\langle f(s), \dot{u}(s)\rangle \leqq \left(\frac{1}{2\rho_0} + \frac{\rho_0}{2}\|\dot{u}(s)\|_0^2\right)\|f(s)\|_0$$

we conclude that

$$\dot{V}^s(v)=\left[\dot{E}^s(v)-\left(\frac{1}{2\rho_0}+E^s(v)\right)\|f(s)\|_0\right]\exp\left(\int_s^\infty \|f(\tau)\|_0\,d\tau\right)\leqq 0, \tag{4.13}$$

where

$$\begin{aligned}\dot{E}^s(v)&\equiv \lim_{\varepsilon\to 0^+}\sup\frac{1}{\varepsilon}[E^{s+\varepsilon}(\omega^s(v,\varepsilon))-E^s(v)]\\ &=-\frac{1}{2}\langle\dot{\rho}(s)\dot{u}(s),\dot{u}(s)\rangle+\frac{1}{2}\langle\dot{A}(s)u(s),u(s)\rangle\\ &\quad+\langle f(s),\dot{u}(s)\rangle-\frac{1}{2}\int_{-\infty}^{s}\langle G_s(s-\tau,s)(u(s)-u(\tau)),u(s)-u(\tau)\rangle\,d\tau.\end{aligned}$$

Since $\mathscr{C}_1$ is dense in $\mathscr{C}_0$ and we have bounds on $u(t)$ by (2.27), the validity of (4.13) is established in the case $v\in\mathscr{C}_0$.

We now establish the equicontinuity of the one-parameter family of maps $V^s(\cdot)\colon \mathscr{C}_0\to\mathbb{R}$, with parameter $s\in\mathbb{R}$, by observing that

$$\begin{aligned}|V^s(v_1)-V^s(v_2)|\leqq\Big\{&\frac{1}{2}|\langle\rho(s)\dot{v}_1(0),\dot{v}_1(0)\rangle-\langle\rho(s)\dot{v}_2(0),\dot{v}_2(0)\rangle|\\ &+\frac{1}{2}|\langle A(s)v_1(0),v_1(0)\rangle-\langle A(s)v_2(0),v_2(0)\rangle|\\ &+\frac{1}{2}\Big|\int_{-\infty}^{s}[\langle G(s-\tau,s)(v_1(0)-v_1(\tau-s)),v_1(0)-v_1(\tau-s)\rangle\\ &-\langle G(s-\tau,s)(v_2(0)-v_2(\tau-s)),v_2(0)-v_2(\tau-s)\rangle]\,d\tau\Big|\Big\}\\ &\cdot\exp\left(\int_{-\infty}^{\infty}\|f(\tau)\|_0\,d\tau\right),\end{aligned} \tag{4.14}$$

$$\begin{aligned}&|\langle\rho(s)\dot{v}_1(0),\dot{v}_1(0)\rangle-\langle\rho(s)\dot{v}_2(0),\dot{v}_2(0)\rangle\\ &\qquad\leqq\|\rho(s)\|_0(\|\dot{v}_1(0)\|_0+\|\dot{v}_2(0)\|_0)|||v_1-v_2|||_{\mathscr{C}_0},\end{aligned} \tag{4.15}$$

$$\begin{aligned}&|\langle A(s)v_1(0),v_1(0)\rangle-\langle A(s)v_2(0),v_2(0)\rangle|\\ &\qquad\leqq\|A(s)\|_{\mathscr{L}(H_1;H_{-1})}(\|v_1(0)\|_1+\|v_2(0)\|_1)|||v_1-v_2|||_{\mathscr{C}_0},\end{aligned} \tag{4.16}$$

$$\begin{aligned}&\Big|\int_{-\infty}^{0}[\langle G(-\xi,s)(v_1(0)-v_1(\xi)),v_1(0)-v_1(\xi)\rangle\\ &\qquad-\langle G(-\xi,s)(v_2(0)-v_2(\xi)),v_2(0)-v_2(\xi)\rangle]\,d\tau\Big|\\ &\leqq 4M(|||v_1|||_{\mathscr{C}_0}+|||v_2|||_{\mathscr{C}_0})|||v_1-v_2|||_{\mathscr{C}_0}.\end{aligned} \tag{4.17}$$

The last inequality follows from (2.10).

It is clear that for every $\varepsilon>0$ there exists a $\delta=\delta(v_1,v_2,\varepsilon)$ such that

$$|V^s(v_1)-V^s(v_2)|<\varepsilon\quad\text{for all } s\in\mathbb{R} \tag{4.18}$$

whenever $|||v_1 - v_2|||_{\mathscr{C}_0} < \delta$. This proves the equicontinuity of $V^s(v)$.

From this property and (4.9) and (4.13) we conclude that $V^s(v)$ is a Lyapunov functional for ω (see [3, Definition 5.1]).

We now state and prove the asymptotic stability theorem:

THEOREM 4.1. *Suppose that for every eigensolution φ_n of the eigenvalue problem*[8]

$$C\varphi - \lambda\rho\varphi = 0 \tag{4.19}$$

there is at least one $\xi_n \in [0, \infty)$ such that

$$\dot{G}(\xi_n)\varphi_n \neq 0.^9 \tag{4.20}$$

Let $u(t)$ be the solution of (1.2), (1.3) *with $v \in \mathscr{C}_0$. Then*

$$u(t) \xrightarrow{H_1} 0, \quad t \to \infty, \tag{4.21}$$

$$\dot{u}(t) \xrightarrow{H_0} 0, \quad t \to \infty. \tag{4.22}$$

Proof. Consider the motion

$$\omega^s(v, \cdot): \mathbb{R}^+ \to \mathscr{C}_0 \tag{4.23}$$

of the process ω which originates at the point (s, v). To prove (4.21) and (4.22) it is sufficient to show that

$$\omega^s(v, \xi) \xrightarrow{\mathscr{C}_0} 0 \quad \text{as } \xi \to \infty. \tag{4.24}$$

For this purpose we first establish a bound for $|||\omega^s(v, \xi)|||_{\mathscr{C}_0}$.

From (4.13) we see that for any $t \geqq s$

$$V^t(\omega^s(v, t-s)) \leqq V^s(v) \tag{4.25}$$

and note that E, as given by (4.8), can be rewritten in the form

$$\begin{aligned} E^t(\omega^s(v, t-s)) &= \frac{1}{2}\langle \rho(t)\dot{u}(t), \dot{u}(t)\rangle + \frac{1}{2}\langle A(t)u(t), u(t)\rangle \\ &\quad -\frac{1}{2}\int_{-\infty}^{t} \langle G(t-\tau, t)(u(t)-u(\tau)), (u(t)-u(\tau))\rangle\, d\tau. \end{aligned} \tag{4.26}$$

From (2.5), (4.4), (4.25), (4.26) and the fact that $l\|u(t)\|_1 \geqq \|u(t)\|_0$ for all $t \in (-\infty, \infty)$, we obtain

$$\sup_{[s,\infty)} \|\dot{u}(t)\|_0 \leqq \left(\frac{2}{\rho_0} V^s(v)\right)^{1/2}, \tag{4.27}$$

$$\sup_{[s,\infty)} \|u(t)\|_1 \leqq \left(\frac{2}{a_0} V^s(v)\right)^{1/2}, \tag{4.28}$$

---

[8] Since $C^{-1} \in \mathscr{L}(H_0; H_2)$ is a positive compact operator on H_0, (4.19) possesses a sequence $\{\lambda_n\}$ of real positive eigenvalues and the corresponding sequence $\{\varphi_n\}$ of eigensolutions is complete in H_0.

[9] Unless $G(\xi)$ is of a very special type, this assumption will be satisfied.

$$\sup_{[s,\infty)} \|u(t)\|_0 \leqq l\left(\frac{2}{a_0} V^s(v)\right)^{1/2}. \tag{4.29}$$

Now we assume temporarily that $v \in \mathscr{B}_0$. This will allow us to establish a bound for $|||\omega^s(v, \xi)|||_{\mathscr{B}_0}$. Let

$$G(1)(t-\tau, t) \equiv \int_\tau^t G_t(t-\xi, t)\, d\xi + G(0, t) + \dot{C}(t), \tag{4.30}$$

$$f_1(t) \equiv -\int_{-\infty}^t G_t(t-\xi, t) u(-\infty)\, d\xi - G(0, t)u(-\infty) - \dot{C}(t)u(-\infty). \tag{4.31}$$

The differentiation of (1.2) will give

$$\frac{d}{dt}\left(\frac{d}{dt}(\rho(t)\dot{u}(t))\right) + C(t)\dot{u}(t) + \int_{-\infty}^t G^{(1)}(t-\tau, t)\dot{u}(\tau)\, d\tau = \dot{f}(t) + f_1(t). \tag{4.32}$$

By combining the term

$$\int_{-\infty}^s G(t-\tau, t)\dot{u}(\tau)\, d\tau \tag{4.33}$$

with $f_1(t)$ we obtain

$$\frac{d}{dt}\left(\frac{d}{dt}(\rho(t)\dot{u}(t))\right) + C(t)\dot{u}(t) + \int_s^t G^{(1)}(t-\tau, t)\dot{u}(\tau)\, d\tau = \dot{f}(t) + f_1^1(t) \tag{4.34}$$

where $f_1^1(t)$ represents the new term after the addition of (4.33) to $f_1(t)$. We define a functional $Q(v)(t)$ by

$$Q(v)(t) \equiv \left[\frac{1}{2\rho_0} + J(v)(t)\right] \exp\left(-\int_s^t (\|\dot{f}(\tau)\|_0 + \|f_1^1(\tau)\|_{-1})\, d\tau\right) \tag{4.35}$$

for $v \in \mathscr{B}_0$, where

$$\begin{aligned} J(v)(t) \equiv {} & \frac{1}{2}\langle \ddot{\rho}(t)\dot{u}(t), \dot{u}(t)\rangle + \frac{1}{2}\langle \rho(t)\ddot{u}(t), \ddot{u}(t)\rangle + \frac{1}{2}\langle A^{(1)}(t)\dot{u}(t), \dot{u}(t)\rangle \\ & - \frac{1}{2}\int_s^t \langle G^{(1)}(t-\tau, t)(\dot{u}(t) - \dot{u}(\tau)), \dot{u}(t) - \dot{u}(\tau)\rangle\, d\tau, \end{aligned} \tag{4.36}$$

$$A^{(1)}(t) \equiv C(t) + \int_s^t G^{(1)}(t-\tau, t)\, d\tau. \tag{4.37}$$

We make the additional assumptions,

$$\langle A^{(1)}(t)w, w\rangle \geqq a_1\|w\|_1 \quad \text{for all } w \in H_1, \quad t \in (-\infty, \infty), \tag{4.38}$$

$$\langle G^{(1)}(\xi, t)w, w\rangle \leqq 0 \quad \text{for all } w \in H_1, \quad \xi \in [0, \infty), \quad t \in (-\infty, \infty), \tag{4.39}$$

$$\langle G_t^{(1)}(t-\tau, t)w, w\rangle \geqq 0 \quad \text{for all } w \in H_1, \quad \tau \in (-\infty, t], \quad t \in (-\infty, \infty), \tag{4.40}$$

$$\langle \dot{A}^{(1)}(t)w, w\rangle \leqq 0 \quad \text{for all } w \in H_1, \quad t \in (-\infty, \infty), \tag{4.41}$$

where

$$\dot{A}^{(1)}(t)=\dot{C}(t)+G^{(1)}(0,t)+\int_s^t G_t^{(1)}(t-\tau,t)\,d\tau. \tag{4.42}$$

With these assumptions we can easily prove the following inequalities:

$$Q(v)(t)\geqq 0, \tag{4.43}$$

$$\frac{d}{dt}Q(v)(t)\leqq 0, \tag{4.44}$$

and hence

$$Q(v)(s)\geqq Q(v)(t)\quad \text{for } t\in[s,\infty). \tag{4.45}$$

On account of (2.5), (4.38) and (4.45) we establish

$$\sup_{[s,\infty)}\|\ddot{u}(t)\|_0\leqq\left[\frac{2}{\rho_0}Q(v)(s)\right]^{1/2}\exp\left(\frac{1}{2}\int_s^\infty(\|\dot{f}(\tau)\|_0+\|f_1^1(\tau)\|_{-1})\,d\tau\right), \tag{4.46}$$

$$\sup_{[s,\infty)}\|\dot{u}(t)\|_1\leqq\left[\frac{2}{a_1}Q(v)(s)\right]^{1/2}\exp\left(\frac{1}{2}\int_s^\infty(\|\dot{f}(\tau)\|_0+\|f_1^1(\tau)\|_{-1})\,d\tau\right). \tag{4.47}$$

From (4.3), (4.46) and remark 2.1 we deduce that $\|u(t)\|_2$ is also uniformly bounded on $[s,\infty)$. Then the motion $\omega^s(v,\cdot)$ of ω through (s,v) is uniformly continuous on $[s,\infty)$ and its orbit is bounded in $\mathscr{B}_0$ and precompact in $\mathscr{C}_0$ by Lemma 2.2. From Proposition 3.2 of [3] it follows that the ω-limit set $\omega^s(v)$ of the motion is nonempty, compact and

$$\omega^s(v,\xi)\xrightarrow{\mathscr{C}_0}\omega^s(v)\quad \text{as } \xi\to\infty. \tag{4.48}$$

Now we are ready to use the invariance principle for compact processes. Actually in our case, the existence of a Lyapunov functional will enable us to use even a more powerful stability theorem which results from the combination of the invariance principle with the Lyapunov functional.

Let $v_1\in\omega^s(v)$. Using Prop. 5.1 of [3] and the fact that $\omega^s(v,\xi)$ is asymptotically a dynamical system with asymptotic hull $\{\bar{\omega}\}$ we deduce that

$$\bar{\omega}(v_1,\xi)\in\omega^s(v)\quad \text{for all } \xi\in\mathbb{R}^+. \tag{4.49}$$

Next we proceed to show that $\omega^s(v)$ is the set $\{0\}$. We observe that $Z\colon\mathscr{C}_0\to\mathbb{R}$ defined by

$$\begin{aligned}Z(v)\equiv{}&\frac{1}{2}\langle\rho\dot{y}(t),\dot{y}(t)\rangle+\frac{1}{2}\left\langle\left(C+\int_0^\infty G(\xi)\,d\xi\right)y(t),y(t)\right\rangle\\&-\frac{1}{2}\int_{-\infty}^t\langle G(t-\tau)(y(t)-y(\tau)),y(t)-y(\tau)\rangle\,d\tau\end{aligned} \tag{4.50}$$

is a Lyapunov functional for the dynamical system $\bar{\omega}$. Here $y(t)$ denotes the solution of (3.17), (3.18). Since $v_1\in\omega^s(v)$, by Prop. 5.1 of [3], we have $\dot{Z}(v_1)=0$. This, in turn, gives

$$\dot{G}(t-\tau)(y(t)-y(\tau))=0,\qquad t\in(-\infty,\infty),\quad \tau\in(-\infty,t), \tag{4.51}$$

This implies that $y(t)=0$ for $y(-\infty, \infty)$ as shown in [5, pp. 305–306]. Thus we have

$$\omega^s(v)=\{0\}. \tag{4.52}$$

The result follows from (4.48).

Now we can remove the assumption $v \in \mathscr{B}_0$ and establish the theorem for $v \in \mathscr{C}_0$ by using an argument similar to the one given in [5, p. 306].

5. Applications to linear viscoelasticity. Let Ω be a bounded domain in E^3. In this section we consider the equations of linear viscoelasticity (1.4), in the cylinder $\Omega\times(-\infty, \infty)$, together with homogeneous boundary conditions

$$\mathbf{u}(\mathbf{x}, t)=\mathbf{0}, \qquad \mathbf{x}\in \partial\Omega, \quad t\in(-\infty, \infty).^{10} \tag{5.1}$$

The history of $\mathbf{u}(\mathbf{x}, t)$ up to a given time s is assigned for the entire body:

$$\mathbf{u}(\mathbf{x}, t)=\mathbf{v}(\mathbf{x}, t), \qquad \mathbf{x}\in\Omega, \quad t\in(-\infty, s]. \tag{5.2}$$

Let $\mathbf{C}_0^\infty(\Omega)$ denote the set of three-dimensional vector fields with compact support in Ω and components in $\mathbf{C}^\infty(\Omega)$. We obtain Hilbert spaces H_0, H_1 and H_2, respectively, by completion of $\mathbf{C}_0^\infty(\Omega)$ under the norms induced by the inner products

$$\langle \mathbf{w}, \mathbf{v}\rangle \equiv \int_\Omega w_i v_i \, dV, \tag{5.3}$$

$$\langle \mathbf{w}, \mathbf{v}\rangle_1 \equiv \int_\Omega \frac{\partial w_i}{\partial x_j}\frac{\partial v_i}{\partial x_j}\, dV, \tag{5.4}$$

$$\langle \mathbf{w}, \mathbf{v}\rangle_2 \equiv \int_\Omega \frac{\partial^2 w_i}{\partial x_j^2}\frac{\partial^2 v_i}{\partial x_j^2}\, dV. \tag{5.5}$$

The space H_{-1} will now be defined as the completion of $\mathbf{C}_0^\infty(\Omega)$ ny means of the norm

$$\|\mathbf{v}\|_{-1} \equiv \sup_{\mathbf{w}\in H_1} \frac{|\langle \mathbf{w}, \mathbf{v}\rangle|}{\|\mathbf{w}\|_1} \tag{5.6}$$

and $\langle\cdot, \cdot\rangle$ is extended onto $H_{-1}\times H_1$ as a continuous bilinear form. Obviously $H_2\subset H_1\subset H_{-1}$ algebraically and topologically. Moreover, by Rellich's theorem, the injection of H_i into H_{i-1}, $i=0, 1, 2$, is compact.

It is clear that the mixed history boundary value problem for (1.4) can be reduced to an abstract history value problem of the type (1.2) and (1.3) for the above selection of H_2, H_1, H_0 and H_{-1}, provided that $\rho(\mathbf{x}, t)$, $C_{ijkl}(\mathbf{x}, t)$ (for fixed t) and $G_{ijkl}(\mathbf{x}, \xi, t)$ (for fixed ξ and t) are Lebesgue measurable functions essentially bounded on Ω. The conditions imposed in §§ 2, 3 and 4 on the operators $\rho(t)$, $C(t)$ and $G(\xi, t)$ can easily be formulated in terms of $\rho(\mathbf{x}, t)$, $C_{ijkl}(\mathbf{x}, t)$ and $G_{ijkl}(\mathbf{x}, \xi, t)$. For example, equations (2.1), (2.2), (2.3), (2.4), (2.5), (2.6) and (2.7) will be

[10] More general classes of homogeneous boundary conditions can be considered alternatively with slight modifications.

satisfied if and only if

$$C_{ijkl}(\mathbf{x}, t) = C_{klij}(\mathbf{x}, t) \quad \text{for } \mathbf{x} \in \Omega \text{ and } t \in (-\infty, \infty), \tag{5.7}$$

$$\int_\Omega C_{ijkl}(\mathbf{x}, t) \frac{\partial w_i}{\partial x_j} \frac{\partial w_k}{\partial x_l} dV \geqq K \int_\Omega \frac{\partial w_i}{\partial x_j} \frac{\partial w_i}{\partial x_j} dV \qquad \text{for all } \mathbf{w} \in H_1, \quad t \in (-\infty, \infty), \tag{5.8}$$

$$G_{ijkl}(\mathbf{x}, \xi, t) = G_{klij}(\mathbf{x}, \xi, t) \quad \text{for } \mathbf{x} \in \Omega, \quad \xi \in [0, \infty) \text{ and } t \in (-\infty, \infty), \tag{5.9}$$

$$\int_\Omega G_{ijkl}(\mathbf{x}, \xi, t) \frac{\partial w_i}{\partial x_j} \frac{\partial w_k}{\partial x_l} dV \geqq 0 \qquad \text{for all } \mathbf{w} \in H_1, \quad \xi \in [0, \infty), \quad t \in (-\infty, \infty), \tag{5.10}$$

$$\rho(\mathbf{x}, t) \geqq \rho_0 > 0 \quad \text{for } \mathbf{x} \in \Omega, \quad t \in (-\infty, \infty), \tag{5.11}$$

$$\frac{\partial \rho(\mathbf{x}, t)}{\partial t} \geqq 0, \quad \mathbf{x} \in \Omega, \quad t \in (-\infty, \infty), \tag{5.12}$$

$$\frac{\partial^2 \rho(\mathbf{x}, t)}{\partial t^2} \geqq 0, \quad \mathbf{x} \in \Omega, \quad t \in (-\infty, \infty). \tag{5.13}$$

Under the conditions (5.7)–(5.13) and some mild smoothness restrictions on the time behavior of $\rho(x, t)$, $C_{ijkl}(x, t)$, $G_{ijkl}(x, \xi, t)$ and $f_i(x, t)$, Theorem 2.3 establishes the existence of a unique solution of equations (1.4), (5.1), (5.2). We note that the solution obtained is not necessarily smooth in $\mathbf{x}$. Additional hypotheses should be made in order to establish the existence of a classical solution. We refer to Edelstein [9] about the necessary hypotheses for classical solutions of quasistatic viscoelasticity equations.

Similarly, the asymptotic behavior of solutions of equations (1.4), (5.1), (5.2) can be investigated through Theorem 4.1. The assumptions required in order to apply this theorem can be easily formulated. For example equations (4.1), (4.2), (4.4), (4.5) will be satisfied, if and only if, for all $\mathbf{w} \in H_1$,

$$\int_\Omega \frac{\partial G_{ijkl}(\mathbf{x}, \xi, t)}{\partial t} \frac{\partial w_i}{\partial x_j} \frac{\partial w_k}{\partial x_l} dV \leqq 0, \qquad \xi \in [0, \infty), \quad t \in (-\infty, \infty), \tag{5.14}$$

$$\int_\Omega \frac{\partial G_{ijkl}(\mathbf{x}, \xi, t)}{\partial \xi} \frac{\partial w_i}{\partial x_j} \frac{\partial w_k}{\partial x_l} dV \leqq 0, \qquad \xi \in [0, \infty), \quad t \in (-\infty, \infty), \tag{5.15}$$

$$\int_\Omega A_{ijkl}(\mathbf{x}, t) \frac{\partial w_i}{\partial x_j} \frac{\partial w_k}{\partial x_l} dV \geqq a_0 \int_\Omega \frac{\partial w_i}{\partial x_l} \frac{\partial w_i}{\partial x_l} dV \qquad \text{for } a_0 > 0, \quad t \in (-\infty, \infty), \tag{5.16}$$

$$\int_\Omega \frac{\partial A_{ijkl}(\mathbf{x}, t)}{\partial t} \frac{\partial w_i}{\partial x_j} \frac{\partial w_k}{\partial x_l} dV \leqq 0, \qquad t \in (-\infty, \infty), \tag{5.17}$$

where

$$A_{ijkl}(\mathbf{x}, t) \equiv C_{ijkl}(\mathbf{x}, t) - \int_{-\infty}^{t} G_{ijkl}(\mathbf{x}, t - \tau, t)\, d\tau. \tag{5.18}$$

The remaining assumptions can be formulated in terms of $\rho(\mathbf{x}, t)$, $C_{ijkl}(\mathbf{x}, t)$ and $G_{ijkl}(\mathbf{x}, \xi, t)$ in a similar way. Under these conditions and assuming that ρ, C_{ijkl} and G_{ijkl} approach time-independent values at large times, Theorem 4.1 proves the asymptotic stability of solutions of (1.4), (5.1), (5.2).

The mechanical interpretation of some of these assumptions is quite clear. For example (5.7) implies a hyperelastic behavior and (5.8) expresses the stability criterion for the instantaneous elastic response of the viscoelastic material. Similarly (5.16) gives the condition for static stability. Equation (5.9) expresses the symmetry of the relaxation function and has been given a mechanistic characterization by Day [7]. Assumptions (5.10) and (5.15) have their origins in experiments with one-dimensional viscoelastic material, for which it was observed that the relaxation function is nonnegative and monotonically decreasing with time. In fact, these properties were interpreted by Day [8] by a characterization in terms of the work done on the material. $\rho(\mathbf{x}, t)$, being the density of the material, is a positive quantity as shown in (5.11). Equations (5.12), (5.13), (5.14) and (5.17) express the assumption that the change in the material response with time is monotone in a certain sense which is compatible with the requirement that the viscoelastic material under consideration approaches a steady state at large times.

To give a mechanical interpretation of (4.3) we consider the one-dimensional viscoelasticity equations:

$$\rho \frac{\partial^2 u(x,t)}{\partial t^2} = c\frac{\partial^2 u(x,y)}{\partial x^2} - \int_{-\infty}^{t} g(t-\tau)\frac{\partial^2 u(x,\tau)}{\partial x^2}\,d\tau + f(x,t), \tag{5.19}$$

where ρ and c are constants and g is independent of x. In this case

$$\|G(\xi, t)\|_{\mathcal{L}(H_2;H_0)} = g(\xi), \tag{5.20}$$

and noting that $c \geqq K$, we observe that (4.3) becomes

$$c - \int_0^\infty g(\xi)\,d\xi > 0. \tag{5.21}$$

This is the static stability condition for one-dimensional viscoelasticity equations, that is, the one-dimensional analogue of (5.16). Although in the multidimensional case, (4.3) is not equivalent to (5.16), it is in the same spirit. By observing that (3.5) is a similar condition to (4.3), we establish its relation to static stability. We also observe that in this case (4.4) reduces to (5.21), hence coincides with (4.3). In the one-dimensional model, the remaining assumptions can be given mechanistic interpretations. In terms of c and $g(t-\tau)$, the inequalities (4.38), (4.39), (4.40) and (4.41) reduce to the one-dimensional analogues of (4.4), (2.4), (4.2) and (2.4), respectively, which have been formulated above to be the inequalities (5.16), (5.10), (5.15) and (5.10).

Finally, we would like to remark about the general character of some of the assumptions made in §§ 2, 3 and 4. It will be remembered that we have interpreted (5.12)–(5.14), (5.17) in terms of the monotonicity of the material response and (3.5), (4.3), (4.4), (4.38)–(4.41) in terms of their one-dimensional analogues. All these assumptions are rather stringent and they were, in fact, necessitated for the proofs of various results in the investigation of the asymptotic stability of the viscoelastic material under consideration. Nevertheless, these assumptions are easily satisfied in the applications. For example, (5.12) and (5.13) are satisfied if the material has a constant density. As yet, there do not seem to be any different mechanical interpretations of the abovementioned assumptions which are based on exact mathematical arguments of the physical phenomena such as the case in [1], [7], [8] and [16]. On the other hand, it is interesting to note that (5.8) has been given a new meaning in [1] after this paper had been first sent for publication. Also another condition in [1, eq. (13)] is in the same sense as (4.39), both of these assumptions being necessary for the asymptotic decay of acceleration waves in the viscoelastic materials of the type considered in this paper. This situation gives reason to hope that similarly some of the hypotheses now needed in this paper will be given a firmer foundation in rational mechanics as the research in this area advances.

Acknowledgment. This paper is based on a dissertation submitted to Cornell University in partial fulfillment of the requirements for the Ph.D. degree. The author is deeply indebted to his thesis advisor, Professor C. M. Dafermos of Brown University, who suggested this topic and supervised it during its preparation.

REFERENCES

[1] F. BLOOM, *On the growth behavior of one-dimensional acceleration waves in the linear theory of viscoelasticity*, J. Appl. Math. Phys., 26 (1975), pp. 361–367.
[2] B. D. COLEMAN AND W. NOLL, *Foundations of linear viscoelasticity*, Rev. Modern Phys., 33 (1961), pp. 239–249.
[3] C. M. DAFERMOS, *An invariance principle for compact processes*, J. Differential Equations, 9 (1971), pp. 239–252.
[4] ———, *An abstract Volterra equation with applications to linear viscoelasticity*, Ibid., 7 (1970), pp. 544–569.
[5] ———, *Asymptotic stability in viscoelasticity*, Arch. Rational Mech. Anal., 37 (1970), pp. 297–308.
[6] ———, *Applications of the invariance principle for compact processes. I. Asymptotically dynamical systems*, J. Differential Equations, 9 (1971), pp. 291–299.
[7] W. A. DAY, *Time-reversal and the symmetry of the relaxation function of a linear viscoelastic material*, Arch. Rational Mech. Anal., 40 (1971), pp. 155–159.
[8] ———, *On monotonicity of the relaxation functions of viscoelastic materials*, Proc. Cambridge Philos. Soc., 67 (1970), pp. 503–508.
[9] W. S. EDELSTEIN, *Existence of solutions to the displacement problem for quasistatic viscoelasticity*, Arch. Rational Mech. Anal., 22 (1966), pp. 121–128.
[10] W. S. EDELSTEIN AND M. E. GURTIN, *Uniqueness theorems in the linear dynamic theory of anisotropic viscoelastic solids*, Ibid., 17 (1964), pp. 47–60.
[11] J. K. HALE, *Dynamical systems and stability*, J. Math. Anal. Appl., 26 (1969), pp. 39–59.
[12] J. J. LEVIN, *The asymptotic behavior of the solution of a Volterra equation*, Proc. Amer. Math. Soc., 14 (1963), pp. 534–541.
[13] R. C. MACCAMY, *Exponential stability for a class of functional differential equations*, Arch. Rational Mech. Anal., 40 (1971), pp. 120–138.
[14] F. ODEH AND I. TADJBAKHSH, *Uniqueness in the theory of viscoelasticity*, Ibid., 18 (1965), pp. 244–250.
[15] V. VOLTERRA, *Sur la théorie mathématique des phénomènes héréditaires*, J. Math. Pures Appl., 7 (1928), pp. 249–298.
[16] C. C. WANG, *The principle of fading memory*, Arch. Rational Mech. Anal., 18 (1965), pp. 343–366.

SIAM J. MATH. ANAL.
Vol. 9, No. 1, February 1978

GREEN'S FUNCTIONS FOR LINEAR SECOND ORDER SYSTEMS*

K. A. HEIMES†

Abstract. A representation is obtained for the Green's function associated with the two point boundary problem $y''=Ay'+By+f$, $y(0)=y(T)=0$, where y and f are vector valued, A and B are closed linear operators. Certain results are applied to solving nonhomogeneous partial differential equations.

1. Introduction. The purpose of this paper is to obtain a representation for the Green's function, or kernel, $G(t, s)$ so that

$$y(t)\equiv\int_0^T G(t,s)f(s)\,ds \tag{1}$$

solves the linear boundary value problem

$$y''(t)=A(t)y'(t)+B(t)y(t)+f(t),\qquad y(0)=y(T)=0. \tag{2}$$

The function f is continuous from the real interval $0\leqq t\leqq T$ into a real or complex Banach space X. A and B will be either (i) constant, closed linear operators defined on a dense subset of X or (ii) strongly continuous from $0\leqq t\leqq T$ into $B(X)$, the algebra of bounded linear operators from X to X, endowed with the operator norm topology. A solution $y(t)$ to (2) will be a twice continuously differentiable function on $0\leqq t\leqq T$ with $y(t)$ in the domain of $B(t)$, $y'(t)$ in the domain of $A(t)$ and (2) is satisfied for $0<t<T$.

When X is the real numbers, $A(t)$ and $B(t)$ are real valued functions so variation of parameters yields the representation

$$\begin{aligned} G(t,s)&=\begin{cases} u(s)[u(t)v(T)-u(T)v(t)]/(u(T)D(s)), & 0\leqq s\leqq t,\\ u(t)[u(s)v(T)-u(T)v(s)]/(u(T)D(s)), & t\leqq s\leqq T,\end{cases}\\ D(s)&\equiv u'(s)v(s)-v'(s)u(s),\end{aligned} \tag{3}$$

in terms of solutions $u(t)$, $v(t)$ to the homogeneous equation $y''=Ay'+By$ with initial conditions $u(0)=v'(0)=0$, $u'(0)=v(0)=1$. We can compute u and v for the case A and B constant to obtain

$$G(t,s)=\frac{1}{2}\left[\exp\frac{(t-s)A}{2}\right][C'(T)]^{-1}\begin{cases} C(t+s-T)-C(T+s-t), & 0\leqq s\leqq T,\\ C(s+t-T)-C(T+t-s), & t\leqq s\leqq T,\end{cases} \tag{4}$$

where $C(t)$ is $\cosh Q^{1/2}t$ or $\cos|Q|^{1/2}t$ or $t^2/2$ according as $Q=\frac{1}{4}A^2+B$ is positive, negative, or zero. (We assume, of course, that $C'(T)\neq 0$).

Our principal result is the generalization of the representation in (4) to the case where A, B are constant, densely defined, closed linear operators on X. Basic assumptations are that $\frac{1}{2}A$ generates a C_0 group and that Q generate a cosine operator function ([2]–[5], [8], [9], [11], [12], [14]). This extends results of Krein and Laptev with $A=0$, the zero operator on X, and $B=P^2$ where the Green's function is expressed in terms of the analytic C_0 semigroup generated by $-P$. See

* Received by the editors December 10, 1974, and in final revised form September 29, 1976.
† Department of Mathematics, Iowa State University, Ames, Iowa 50010.

[10, pp. 247–270]. We also give a representation for $G(t, s)$ as a series in powers of A and B with scalar valued coefficients; convergence places strong demands on the forcing function $f(t)$.

Section 2 concerns nonconstant bounded coefficient operators. A representation analogous to (3) is obtained.

We avoid reducing the second order equation in (2) to a first order system as one finds in Conti [1] and Reid [13].

2. Bounded coefficient operators. We assume throughout this section that $A(t)$ and $B(t)$ are strongly continuous from $0 \leqq t \leqq T$ into $B(X)$. An operator P in $B(X)$ is invertible or regular in case P is one-one and onto with P^{-1} in $B(X)$.

It follows from the Banach–Steinhaus theorem that $\|A(t)\| \leqq m$, $\|B(t)\| \leqq m$ on $0 \leqq t \leqq T$ for some constant $m > 0$ and that $A(t)W(t)$, $B(t)W(t)$ are norm continuous from $0 \leqq t \leqq T$ into $B(X)$ when $W(t)$ has that property. Consequently, the Picard iteration scheme shows that the Cauchy problem for

$$Y''(t) = A(t)Y'(t) + B(t)Y(t) \tag{5}$$

in $B(X)$ has a unique solution which is twice continuously differentiable in norm and exists on $0 \leqq t \leqq T$.

THEOREM 1. *For strongly continuous bounded operators $A(t)$, $B(t)$ on $0 \leqq t \leqq T$, let $U(t)$ and $V(t)$ solve* (5) *with*

$$U(0) = V'(0) = 0 = \textit{zero operator}$$

$$U'(0) = V(0) = I = \textit{identity operator.}$$

If $U(t)$ is invertible on $0 < t \leqq T$ then (1) *solves* (2) *for every continuous function f where*

$$G(t, s) = \begin{cases} 0, & 0 = s \leqq t, \\ U(t)[R(T) - R(t)]H^{-1}(s), & 0 < s \leqq t, \\ U(t)[R(T) - R(s)]H^{-1}(s), & t \leqq s \leqq T, \end{cases}$$

$$R(t) = U^{-1}(t)V(t), \qquad H(t) = -U(t)R'(t) \tag{6}$$

Proof. Since $U(t)$ is differentiable and $U^{-1}(t)$ exists, $U^{-1}(t)$ is differentiable on $0 < t \leqq T$. Thus $R(t)$ is defined and differentiable on $0 < t \leqq T$.

From the initial conditions on U, we have

$$\lim_{t \to 0^+} \left\| \frac{1}{t} U(t) - I \right\| = \lim_{t \to 0^+} \left\| \frac{1}{t} \int_0^t [U'(r) - I] \, dr \right\| = 0.$$

Since the inverse map is continuous on the invertible operators of $B(X)$, $\lim_{t \to 0^+} tU^{-1}(t) = I$. Thus $\lim_{t \to 0^+} tH(t) = I$. Now the invertible operators form an open set in $B(X)$ so $tH(t)$ and $H(t)$ are invertible on some interval $0 < t < \delta$. But $H(t)$ solves the differential equation $H' = (A - U'U^{-1})H$ on $0 < t \leqq T$ so H is invertible on $0 < t \leqq T$. In fact, $\lim_{t \to 0^+} (1/t)H^{-1}(t) = I$ so H^{-1} is continuous on $0 \leq t \leq T$ by taking $H^{-1}(0) = 0$.

We have established thus far that $G(t, s)$ is norm continuous from the square $[0, T] \times [0, T]$ into $B(X)$. Clearly $G(0, s) = G(T, s) = 0$. Since $U(t)R(t) = V(t)$, we see that, for fixed s, $G(t, s)$ is a linear combination of $U(t)$ and $V(t)$ for $0 < t < s$

and $s<t<T$. Consequently, $G(t, s)$ is class C^2 and solves (5) as a function of t for fixed s in the indicated regions. Note also that

$$\frac{\partial G}{\partial t}(t+h, s)-\frac{\partial G}{\partial t}(t-k, s)=[U'(t+h)-U'(t-k)]R(T)H^{-1}(s)$$
$$+H(t+h)H^{-1}(s)+[U'(t+h)R(t+h)$$
$$-U'(t-k)R(s)]H^{-1}(s)$$

for $t-k<s<t+h$. This expression has limit I as $h, k\to 0$. Thus $\lim_{h\to 0_+}(1/h)\int_t^{t+h}[(\partial G/\partial t)(t+h, s)-(\partial G/\partial t)(t, s)]f(s)\,ds=f(t)$. Using these facts, one now shows by direct calculation that the function $y(t)$ defined in (1) does indeed solve (2).

COROLLARY. *When the independent solutions $U(t)$, $V(t)$ to equation* (5) *are related by $U(t)=tV(t)$, the Green's function in* (6) *reduces to*

$$G(t, s)=\begin{cases}\dfrac{s(t-T)}{T}V(t)V^{-1}(s), & 0\leqq s\leqq t,\\[2ex] \dfrac{t(s-T)}{T}V(t)V^{-1}(s), & t\leqq s\leqq T.\end{cases} \tag{7}$$

Example. Let $X=R^2$,

$$A(t)=4t\begin{bmatrix}-t^2 & 1\\ 1-t^4 & t^2\end{bmatrix},\qquad B(t)=\begin{bmatrix}-10t^2 & 2\\ 2-10t^4 & 2t^2\end{bmatrix}.$$

Then independent solutions to (5) are

$$V(t)=\begin{bmatrix}1 & t^2\\ t^2 & 1+t^4\end{bmatrix}\quad\text{and}\quad U(t)=tV(t).$$

The solution to the two dimensional system (2) is then easily computed for any forcing vector $f(t)$ using $G(t, s)$ from (7) in (1).

Notice that we can generate a class of problems whose Green's matrix is given by the above corollary. Take any C^2 matrix $V(t)$ with $V(0)=I$ and $V'(0)=0$. Then $V(t)$ and $U(t)=tV(t)$ are fundamental solutions to (5) when $A=2V'V^{-1}$ and $B=V''V^{-1}-\frac{1}{2}A^2$.

3. Unbounded coefficients. Henceforth we use the notation $\mathrm{D}(Q)$ for the domain of the operator Q.

DEFINITION. Let Q be a closed linear operator whose domain $\mathrm{D}(Q)$ is dense in X. We say that Q generates a "cosine operator function" in case there is a strongly continuous function $C(t)$ from $-\infty<t<\infty$ to $B(X)$ with

(i) $C(t)x\in \mathrm{D}(Q)$ for every $x\in \mathrm{D}(Q)$ and all t,

(ii) $x\in \mathrm{D}(Q)$ if and only if $C(t)x$ is twice continuously differentiable in norm with $C''(t)x=C(t)Qx=QC(t)x$ on $-\infty<t<\infty$ and with $C(0)x=x$, $C'(0)x=0$.

Da Prato and Giusti [2], Fattorini [3], and Sova [14] have given conditions on the resolvent of Q which are necessary and sufficient in order that Q generate a cosine function. Their result is analogous to the Hille–Yosida generation theorem for semigroups of operators.

Starting with this definition one can show that the cosine operator satisfies $C(t+s)+C(t-s)=2C(t)C(s)$, $C(0)=I$, $C(t)=C(-t)$, and $C(t)C(s)=C(s)C(t)$ for all real s, t. (cf. [9, p. 95]). Moreover, the bounded operator $S(t)$ defined by $S(t)x=\int_0^t C(s)x\,ds$ is continuously differentiable in t for all $x\in X$, twice continuously differentiable for $x\in \mathrm{D}(Q)$, commutes with $C(t)$, commutes with Q on $\mathrm{D}(Q)$ and solves

$$S''(t)x=QS(t)x=S(t)Qx, \qquad S(0)x=0, \quad S'(0)x=x$$

for fixed $x\in \mathrm{D}(Q)$. Notice that for $x\in \mathrm{D}(Q)$,

$$C'(T)x=\int_0^T C''(s)x\,ds=\int_0^T QC(s)x\,ds=QS(T)x=S(T)Qx.$$

We use this in our first generalization of equation (4).

THEOREM 2. *Let Q generate a cosine function $C(t)$ and let $S(T)=\int_0^T C(s)\,ds$. Assume each of Q and $S(T)$ has a bounded inverse. Then for every continuous function $F(t)$ from $0\leqq t\leqq T$ into X, the function*

$$y(t)=\int_0^T G(t,s)F(s)\,ds \tag{8}$$

solves

$$y''(t)=Qy(t)+F(t), \qquad y(0)=y(T)=0 \tag{9}$$

where

$$G(t,s)\equiv\begin{cases}\frac{1}{2}[C(s+t-T)-C(s+T-t)]Q^{-1}S^{-1}(T), & 0\leqq s\leqq t,\\ \frac{1}{2}[C(t+s-T)-C(t+T-s)]Q^{-1}S^{-1}(T), & t\leqq s\leqq T.\end{cases} \tag{10}$$

Proof. Put $k(s)=Q^{-1}S^{-1}(T)F(s)$. Since $C(t)$ is strongly class C^2 on $\mathrm{D}(Q)$ and $F(s)$ is continuous on $0\leqq s\leqq T$, it follows that $k(s)$ is continuous on $0\leqq s\leqq T$ to $\mathrm{D}(Q)$ and $C(s+t-T)k(s)$ is class C^2 in t uniformly in s on $0\leqq s\leqq T$. Thus, differentiating (8) gives

$$\begin{aligned}y'(t)=&\frac{1}{2}\int_0^t [C'(s+t-T)+C'(s+T-t)]k(s)\,ds\\&+\frac{1}{2}\int_t^T [C'(t+s-T)-C'(t+T-s)]k(s)\,ds\end{aligned} \tag{11}$$

and

$$\begin{aligned}y''(t)&=C'(T)k(t)+\int_0^T QG(t,s)F(s)\,ds\\&=F(t)+QY(t).\end{aligned}$$

The end conditions are easily checked.

THEOREM 3. *Let A, B be constant closed linear operators densely defined on X with $D(Q)=D(B)\subset D(A^2)$ where $Q=\frac{1}{4}A^2+B$. Assume that $\frac{1}{2}A$ generates a C_0 group $U(t)$, that $BU(t)x=U(t)Bx$ for $x\in D(B)$, that Q generates a cosine function $C(t)$ and that Q^{-1} and $S^{-1}(T)$ exists. Then for every continuously differentiable function $f(t)$ on $0\leqq t\leqq T$ into $D(A)$, equation* (2) *has a solution given by*

$$y(t)=U(t)\int_0^T G(t,s)U(-s)f(s)\,ds$$

where $G(t,s)$ is as in (10).

Proof. Let $w(t)$ solve (9) with $F(t)=U(-t)f(t)$. Put $y(t)=U(t)w(t)$. Now $w(t)$ is class C^2 into $D(Q)\subset D(A^2)$ so $y(t)$ is at least class C^1 with

$$y'(t)=U'(t)w(t)+U(t)w'(t). \tag{12}$$

The first term on the right in (12) is again differentiable because $w(t)\in D(A^2)$. Integrating (11) by parts ($w'(t)$ is given by (11)) shows that $w'(t)$ maps $0\leqq t\leqq T$ into $D(Q)\subset D(A)$ so that $U(t)w'(t)$ is in $D(A)$ and differentiable. Thus $y'(t)$ is in $D(A)$ and differentiable with

$$\begin{aligned} y''&=U''w+U'w'+U'w'+Uw'' \\ &=\tfrac{1}{4}A^2Uw+AUw'+U(t)[Qw(t)+U(-t)f(t)]. \end{aligned}$$

From (12) we get $AUw'=Ay'-AU'w=Ay'-\frac{1}{2}A^2Uw$ and this substitution above gives $y''=Ay'+By+f$ where we have also used $UQ=QU$ since U commutes with B. Q.E.D.

If $\frac{1}{2}A$ generates a C_0 group and if B is closed with $D(B)\supset D(A)$, then necessarily Q generates a cosine function [5, p. 250]. In this case the conclusion of Theorem 3 is valid if either f is continuous with value in $D(A^2)$ or if f is continuously differentiable with value in X.

When $D(A)=X$ in Theorem 3, it suffices to have f continuous with value in X.

THEOREM 4. *Let Q generate a cosine function $C(t)$ and put $S(t)=\int_0^t C(s)\,ds$. Assume $S^{-1}(T)$ exists. If $F(t)$ is continuous on $0\leqq t\leqq T$ into $D(Q)$, then* (9) *has a solution given by* (8) *with*

$$G(t,s)\equiv\begin{cases} S(t-T)S(s)S^{-1}(T), & 0\leqq s\leqq t,\\ S(s-T)S(t)S^{-1}(T), & t\leqq s\leqq T.\end{cases} \tag{13}$$

Proof. Since $S'(t)x=C(t)x$ for all $x\in X$,

$$y(t)=\int_0^t S(t-T)S(s)S^{-1}(T)F(s)\,ds+\int_t^T S(t)S(s-T)S^{-1}(T)F(s)\,ds$$

is clearly class C^1 with

$$y'(t)=\int_0^t C(t-T)S(s)S^{-1}(T)F(s)\,ds+\int_t^T C(t)S(s-T)S^{-1}(T)F(s)\,ds.$$

Now $D(Q)$ is invariant under both the cosine operator $C(t)$ and the sine operator

$S(t)$ so with $F(s)$ continuous into $D(Q)$ we have

$$y''(t)=[C(t-T)S(t)-C(t)S(t-T)]S^{-1}(T)F(t)$$
$$+\int_0^t C'(t-T)S(s)S^{-1}(T)F(s)\,ds+\int_t^T C'(t)S(s-T)S^{-1}(T)F(s)\,ds.$$

Setting $P(t)=C(t-T)S(t)-C(t)S(t-T)$ gives $P(T)=S(T)$ and $P'(t)x=0$ for all $x\in D(Q)$ so that $P(t)=S(T)$ on $0\leqq t\leqq T$. Also $C'(t)x=QS(t)x$ for $x\in D(Q)$. Thus the above expression reduces to $y''(t)=F(t)+Qy(t)$. Q.E.D.

The requirement that Q^{-1} and/or $S^{-1}(T)$ exist seems difficult to verify in examples. For a particular forcing function $F(t)$, however, it is clear that the theorems and formulas are correct provided there is a function $g(t)$ into $D(Q)$ so that $QS(T)g(t)=F(t)$ (or $S(T)g(t)$ in Theorem 4). Moreover, by the cosine function generation theorem [2, p. 358], Q^{-1} automatically exists as a bounded operator on X if we replace Q by $Q-\lambda I$ for $\lambda>0$ sufficiently large. This simply involves perturbing B by a scalar.

Example. Let $Q=d^2/dx^2$ and let X denote the uniformity continuous bounded functions on $-\infty<x<\infty$ with supremum norm. Q generates the cosine operator $C(t)$ defined by $C(t)\varphi(x)=\frac{1}{2}[\varphi(x+t)+\varphi(x-t)]$. Let $g(x,s)\in D(Q)$ for $0\leqq s\leqq T$. Replacing $Q^{-1}S^{-1}(T)F(s)$ in (8) of Theorem 2 by $g(x,s)$ we have that

$$\begin{aligned}U(t,x)=\frac{1}{4}\int_0^t &[g(x+s+t-T,s)+g(x-s-t+T,s)\\&-g(x+s+T-t,s)-g(x-s-T+t,s)]\,ds\\+\frac{1}{4}\int_t^T &[g(x+s+t-T,s)+g(x-s-t+T,s)\\&-g(x+t+T-s,s)-g(x-t-T+s,s)]\,ds\end{aligned}\tag{14}$$

solves

$$\frac{\partial^2 u}{\partial t^2}=\frac{\partial^2 u}{\partial x^2}+f(x,t),$$
$$u(0,x)=u(T,x)=0\quad\text{for all } x$$

where $f(x,t)=\frac{1}{2}(\partial/\partial x)\{g(x+T,t)-g(x-T,t)\}$. (The variables t and x are interchanged from their usual role in the wave equation).

Equation (14) is good for generating explicit solutions for a reasonable class of problems.

Our final result gives a series representation for the solution to (2).

For integers $n\geqq 1$ define the functions

$$h_n(t,s)=\frac{1}{2T}\begin{cases}[(s+t-T)^{2n}-(s-t+T)^{2n}]/(2n)!, & 0\leqq s\leqq t,\\ [(t+s-T)^{2n}-(t-s+T)^{2n}]/(2n)!, & t\leqq s\leqq T,\end{cases}$$

and

$$g_n(t,s)=\begin{cases} h_1(t,s), & n=1,\\ h_n(t,s)-\sum_{j=1}^{n-1} g_{n-j}(t,s)T^{2j}/(2j+1)!, & n>1.\end{cases}$$

Note that all these functions vanish at $t=0$ and $t=T$.

One verifies directly that $(\partial^2 h_n/\partial t^2)(t,s)=h_{n-1}(t,s)$ for $t<s$ or $t>s$ and, by induction, we find that $(\partial^2 g_n/\partial t^2)(t,s)=g_{n-1}(t,s)$ in the same regions.

Since

$$\lim_{s\to t^-}\frac{\partial h_n}{\partial t}(t,s)-\lim_{s\to t^+}\frac{\partial h_n}{\partial t}(t,s)=T^{2(n-1)}/(2n-1)!$$

for $n\geqq 1$, we obtain inductively

$$\lim_{s\to t^-}\frac{\partial g_n}{\partial t}(t,s)-\lim_{s\to t^+}\frac{\partial g_n}{\partial t}(t,s)=\begin{cases}1, & n=1,\\ 0, & n>1.\end{cases}$$

With these facts, the following lemma is easily verified.

LEMMA. *Let $\frac{1}{2}A$ generate a C_0 group $U(t)$ and let $f(t)$ be continuous on $0\leqq t\leqq T$ into $D(A^2)$. Put $y_0(t)=f(t)$ and $y_n(t)=U(t)\int_0^T g_n(t,s)U(-s)f(s)\,ds$.*

Then for $n\geqq 1$, $y_n(t)$ solves

$$y_n''=Ay_n'-\tfrac{1}{4}A^2y_n+y_{n-1},\qquad y_n(0)=y_n(T)=0.$$

THEOREM 5. *Let $Q=\frac{1}{4}A^2+B$ where A and B commute on a dense subset of X and $\frac{1}{2}A$ generates a C_0 group. With $f(t)$, $y_n(t)$ as in the lemma, suppose $y(t)=\sum_{n=1}^\infty Q^{n-1}y_n(t)$ is class C^2 from $0\leqq t\leqq T$ into X. Then $y(t)$ solves (2).*

Proof.

$$\begin{aligned} y''&=\sum_{n=1}^\infty Q^{n-1}[Ay_n'-\tfrac{1}{4}A^2y_n+y_{n-1}]\\ &=Ay'-\tfrac{1}{4}A^2y+\sum_{k=0}^\infty Q^ky_k\\ &=Ay'-\tfrac{1}{4}A^2y+f+Q\sum_{n=1}^\infty Q^{n-1}y_n\\ &=Ay'+(Q-\tfrac{1}{4}A^2)y+f,\end{aligned}$$

where we have used the convention $Q^0=I$ and the fact that A commutes with Q^n for every $n\geqq 0$.

Although this result is formal, it offers a prospect of computability in certain cases. For example, consider

$$\frac{\partial^2 u}{\partial t^2}+\frac{\partial^2 u}{\partial x^2}=2\frac{\partial^2 u}{\partial x\partial t}+f(t,x)$$

on the rectangle $0\leqq t\leqq T$, $a<x<b$ with $u(0,x)=u(T,x)=0$. If we take $A=2(\partial/\partial x)$ and $B=-\partial^2/\partial x^2 Q=0$ so $u(t,x)=\int_0^T g_1(t,s)f(s,x+t-s)\,ds$ is a solution when $f(t,x)$ is continuous in t and class C^2 in x.

Examples can also be constructed with differential operators where $Q^{n-1}y_n(t)=0$ for all $n \geqq N$. Polynomial forcing functions fall in this class and the formal series then provides a means of obtaining the exact solution.

REFERENCES

[1] R. CONTI, *Recent trends in the theory of boundary value problems for ordinary differential equations*, Boll. Un. Mat. Ital., 22 (1967), pp. 135–178.

[2] G. DA PRATO AND E. GIUSTI, *Una caratterizzazione die generatori di funzioni coseno astratte*, Ibid., 22 (1967), pp. 357–362.

[3] H. O. FATTORINI, *On differential equations in lineae topological spaces, I & II*, J. Differential Equations, 5 (1968), pp. 72–105; 6 (1969), pp. 50–70.

[4] J. A. GOLDSTEIN, *Semigroups and second order differential equations*, J. Functional Analysis, 4 (1969), pp. 50–70.

[5] ———, *On a connection between first and second order differential equations in Banach space*, J. Math. Anal. Appl., 30 (1970), pp. 246–251.

[6] K. A. HEIMES, *Two point boundary problems in Banach space*, J. Differential Equations, 5 (1969), pp. 215–225.

[7] E. HILLE AND R. S. PHILLIPS, *Functional Analysis and Semi-groups*, Colloq. Publ., Vol. 31, American Mathematical Society, Providence, R.I., 1957.

[8a] J. KISYŃSKI, *On operator valued solutions of d'Alemberts functional equation, I*, Colloq. Math., 23 (1971), pp. 107–114.

[8b] ———, *On operator valued solutions of d'Alemberts functional equation II*, Studia Math., 42 (1972), pp. 43–66.

[9] ———, *On cosine operator functions and one-parameter groups of operators*, Studia Math., 44 (1972), pp. 93–105.

[10] S. G. KREIN, *Linear Differential Equations in Banach Space*, AMS Transl., Vol. 29, American Mathematical Society, Providence, RI, 1971.

[11] S. KUREPA, *A cosine functional equation in Hilbert space*, Canad. J. Math., 12 (1960), pp. 45–60.

[12] ———, *A cosine functional equation in Banach algebras*, Acta. Sci. Math. (Szeged), 23 (1962), pp. 255–267.

[13] W. T. REID, *Ordinary Differential Equations*, John Wiley, New York, 1971.

[14) M. SOVA, *Cosine operator functions*, Dissertations Math. Rozprawy Mat., 49 (1966), pp. 1–47.

SIAM J. MATH. ANAL.
Vol. 9, No. 2. April 1978
Copyright © 1978 Society for Industrial and Applied Mathematics

THE ASYMPTOTIC SOLUTION OF A CLASS OF SINGULARLY PERTURBED NONLINEAR BOUNDARY VALUE PROBLEMS VIA DIFFERENTIAL INEQUALITIES*

F. A. HOWES†

Respectfully Dedicated to Professor M. Nagumo on the Occasion of his Seventieth Birthday

Abstract. The singular perturbation problem $\varepsilon y''=f(t, y, y')$, $0<t<1$, $y(0, \varepsilon)$, $y(1, \varepsilon)$ given, is studied under the principal assumption that $f_{y'y'}$ is never zero in the domain of interest. Solutions are shown to exhibit essentially two types of asymptotic behavior: (i) boundary layer behavior and/or (ii) smooth transition from one stable reduced root to another. In addition, an algorithm for the exact determination of conditions guaranteeing such behavior as well as several illustrative examples are discussed. The results are established by applying an extension of the classical Nagumo theory of differential inequalities.

1. Introduction. In this paper we consider the existence and the qualitative behavior of solutions of the singularly perturbed boundary value problem

$$\varepsilon y''=f(t, y, y'), \qquad 0<t<1, \tag{1.1}$$

$$y(0, \varepsilon)=A, \qquad y(1, \varepsilon)=B, \tag{1.2}$$

for small, positive values of the parameter ε. The principal assumptions regarding the function f are that $f=O(|y'|^2)$, as $|y'|\to\infty$, and that the partial derivative $f_{y'y'}$ is never zero in the region of interest. Under additional smoothness and stability restrictions on f, we are able to discuss the asymptotic behavior as $\varepsilon\to 0^+$ of solutions of (1.1), (1.2) for all values of the boundary conditions A and B. In addition, we present an algorithm for the determination of such behavior which follows in a straightforward manner from the sign restrictions on $f_{y'y'}$ and the properties of solutions of appropriate reduced problems. This algorithm will be seen to apply to related problems not explicitly treated here as well as to other classes of nonlinear second-order singularly perturbed boundary value problems.

Although there are several ways of proving the theorems presented below, we choose to employ a method involving the use of differential inequalities which was first introduced by N. I. Briš [2]. Using a result of M. Nagumo [15], Briš was able to study the existence and the asymptotic behavior of solutions of (1.1), (1.2) under the principal assumptions that $f_{y'y'}=O(1)$, as $|y'|\to\infty$, and that $f_{y'}\leqq -k<0$ in a suitable tube around the solution u of the reduced problem, $0=f(t, u, u')$, $u(1)=B$. In essence, he showed that for all sufficiently small values of $\varepsilon>0$, there exists a solution $y=y(t, \varepsilon)$ of (1.1), (1.2) satisfying $y=u+O(|A-u(0)|\exp[-kt\varepsilon^{-1}])+O(\varepsilon)$, for t in $[0, 1]$.

In the present paper we wish to modify these two basic assumptions in such a way as to study a related class of problems. In so doing we will touch upon the work

* Received by the editors January 28, 1976, and in revised form August 5, 1976.

† School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455. This research was conducted while the author was a Visiting Member, Courant Institute of Mathematical Sciences, New York University, under National Science Foundation Grant NSF-GP-37069X. The final revision was performed at University of Wisconsin—Madison under National Science Foundation Grant MCS 76-05979.

of several authors in the area of singularly perturbed boundary value problems. Most notably, the results of S. Haber and N. Levinson [6] will be employed frequently in the discussion of solutions of (1.1), (1.2) for certain values of A and B. Their theorem, which also applies to nonlinear functions f more general than those considered here, becomes even more natural when placed in the context of the algorithm to be outlined below. Later generalizations of this basic result may be found in the survey article of A. B. Vasil'eva [19, Chap. 2] and in the paper of R. E. O'Malley [16]. This article of Vasil'eva also contains an account of the major work done on nonlinear problems of the form (1.1), (1.2) up to 1963.

Together with the previously mentioned paper of Briš [2], the recent paper of F. Dorr, S. Parter and L. Shampine [4, § 5] serves as the motivation for our study of the problem (1.1), (1.2) under the assumption that $f=O(|y'|^2)$, as $|y'|\to\infty$. Specifically, these authors considered the following special case of (1.1), (1.2):

$$\text{(1.3)} \qquad \varepsilon y''=p(t,y)(y')^2+g(t,y)y'-b(t,y)y+F(t,y), \qquad 0<t<1,$$

$$\text{(1.4)} \qquad y(0,\varepsilon)=A, \qquad y(1,\varepsilon)=B,$$

where $p(t,y)\leqq -p_0<0$ and $A<B$, $B>0$. Using a priori estimates derived from maximum principle arguments, they were able to prove the existence of solutions of (1.3), (1.4) for all $\varepsilon>0$, under suitable restructions on the functions g, b and F. However, based solely on the estimates given in [4, § 5], it is less clear how one should proceed to study the asymptotic behavior of these solutions as $\varepsilon\to 0^+$. In the course of the present treatment we will indicate an approach to the study of the asymptotic behavior of solutions of the general problem (1.1), (1.2) which is conceptually simpler than that given by Dorr, Parter and Shampine for (1.3), (1.4) and which is more readily verifiable in concrete applications of the theory.

Other interesting discussions of nonlinear problems of the form (1.1), (1.2) can be found in the paper of W. Wasow [21] and the book of O'Malley [17, Chap. 5]. This book and that of Wasow [20, Chap. 10] contain a wealth of information on the general theory of singular perturbations as well as extensive references to the mathematical and scientific literature. More recently, D. Cohen [3] has studied initial and boundary value problems involving the nonlinear equation (1.1) and he has found solutions which possess unexpected oscillatory properties.

Before considering the various phenomena which solutions of (1.1), (1.2) can display, we discuss in the next two sections the mathematical theory of differential inequalities basic to our approach and also the properties of solutions of reduced problems which play a central role in all that follows.

2. Mathematical preliminaries. The main mathematical tool we employ is a theorem on differential inequalities which was first proved by Nagumo [15] and later refined by L. K. Jackson [13, § 7]. In the context of problems of the form (1.1), (1.2), it may be stated as follows.

THEOREM 2.1. *Assume that the function $f=f(t,y,y')$ is continuous on $[0,1]\times\mathbb{R}^2$ and grows no faster than $(y')^2$, as $|y'|\to\infty$, for t in $[0,1]$ and y bounded, i.e., $f=O(|y'|^2)$, as $|y'|\to\infty$. Assume also that there exist functions α and β of class $C^{(2)}[0,1]$ which satisfy the following inequalities*:

$$\alpha\leqq\beta, \quad \alpha(0,\varepsilon)\leqq A\leqq\beta(0,\varepsilon), \quad \alpha(1,\varepsilon)\leqq B\leqq\beta(1,\varepsilon),$$

and for t in $(0, 1)$,

$$\varepsilon\alpha'' \geqq f(t, \alpha, \alpha'), \qquad \varepsilon\beta'' \leqq f(t, \beta, \beta').$$

Then for all such $\varepsilon > 0$, *the boundary value problem* (1.1), (1.2),

$$\varepsilon y'' = f(t, y, y'), \qquad 0 < t < 1,$$
$$y(0, \varepsilon) = A, \qquad y(1, \varepsilon) = B,$$

has a solution $y = y(t, \varepsilon)$ *with* $\alpha(t, \varepsilon) \leqq y(t, \varepsilon) \leqq \beta(t, \varepsilon)$, *for t in* $[0, 1]$.

This theorem, in the form originally stated by Nagumo [15], was the basis of Briš's treatment of the problem (1.1), (1.2). For more recent applications, see [1], [7] and [8]. In the present discussion of (1.1), (1.2) we will likewise make extensive use of this theorem; however, it will also be necessary to use the following extension of Theorem 2.1 due to P. Habets and M. Laloy [7].

THEOREM 2.2. *Make the same assumptions as in Theorem* 2.1 *with the exception that the functions* α *and* β *are assumed to be continuous and piecewise-*$C^{(2)}$ *on* $[0, 1]$, *i.e., there is a finite partition of* $[0, 1]$, $0 = t_0 < t_1 < \cdots < t_n = 1$, *and in each subinterval* (t_i, t_{i+1}), α *and* β *are of class* $C^{(2)}$. *Furthermore, for t in* (t_i, t_{i+1}),

$$\varepsilon\alpha'' \geqq f(t, \alpha, \alpha'), \qquad \varepsilon\beta'' \leqq f(t, \beta, \beta');$$

and at the partition points t_i, $D_l\alpha(t_i) \leqq D_r\alpha(t_i)$, $D_l\beta(t_i) \geqq D_r\beta(t_i)$, *where* D_l, D_r *denote the lefthand, respectively, righthand derivative. Then the conclusion of Theorem* 2.1 *is valid*

The bounding solutions α and β are thus allowed to have finitely many "corners", with the stipulation that the one-sided derivatives at the corner points satisfy the correct inequalities. Theorem 2.2 will be used in studying solutions of the problem (1.1), (1.2) which possess interior nonuniformities in their derivatives.

Since the function f appearing in (1.1) is assumed to satisfy $f = O(|y'|^2)$, as $|y'| \to \infty$, our study of the existence and the asymptotic behavior of solutions of (1.1), (1.2) for small $\varepsilon > 0$ is reduced via Theorems 2.1 and 2.2 to the construction of sufficiently accurate bounding functions α and β. In verifying that such α, β (to be given below) do in fact satisfy the correct differential inequalities, we require the following form of Taylor's theorem for expanding the function f around a given function u. In particular, let $\sigma = \sigma(t, \varepsilon)$ and $u = u(t)$ be given, differentiable functions; then

$$\begin{aligned} f(t, \sigma, \sigma') &= f(t, u, u') + \{f(t, \sigma, u') - f(t, u, u')\} + \{f(t, \sigma, \sigma') - f(t, \sigma, u')\} \\ &= f(t, u, u') + \sum_{j=1}^{N-1} \frac{1}{j!} \partial_y^j f(t, u, u')(\sigma - u)^j + \frac{1}{N!} \partial_y^N f(t, u + \mathcal{O}_1(\sigma - u), u')(\sigma - u)^N \\ &\quad + \sum_{j=1}^{M-1} \frac{1}{j!} \partial_{y'}^j f(t, \sigma, u')(\sigma' - u')^j + \frac{1}{M!} \partial_{y'}^M f(t, \sigma, u' + \mathcal{O}_2(\sigma' - u'))(\sigma' - u')^M, \end{aligned}$$

where $0 < \mathcal{O}_1, \mathcal{O}_2 < 1$ and $N, M \geqq 1$. We will be interested in the case $M = 2$; and as will become clear later, we require the following further expansion of $f_{y'}(t, \sigma, u')$:

$$f_{y'}(t, \sigma, u') = f_{y'}(t, u + (\sigma - u), u')$$
$$= f_{y'}(t, u, u') + f_{yy'}(t, u + \mathcal{O}_3(\sigma - u), u')(\sigma - u).$$

Thus the final expansion takes the form

$$\begin{aligned} f(t, \sigma, \sigma') = f(t, u, u') &+ \sum_{j=1}^{N-1} \frac{1}{j!} \partial_y^j f(t, u, u')(\sigma - u)^j \\ &+ \frac{1}{N!} \partial_y^N f(t, u + \mathcal{O}_1(\sigma - u), u')(\sigma - u)^N \\ &+ \{f_{y'}(t, u, u') + f_{yy'}(t, u + \mathcal{O}_3(\sigma - u), u')(\sigma - u)\}(\sigma' - u') \\ &+ \frac{1}{2} f_{y'y'}(t, \sigma, u' + \mathcal{O}_2(\sigma' - u'))(\sigma' - u')^2, \end{aligned}$$

for $0 < \mathcal{O}_1, \mathcal{O}_2, \mathcal{O}_3 < 1$ and $N \geqq 1$.

3. The reduced problem. The starting point for our investigation of the nonlinear problem (1.1), (1.2) is the study of the reduced equation obtained by formally setting $\varepsilon = 0$ in equation (1.1), i.e.,

$$(3.1) \qquad 0 = f(t, u, u').$$

In what follows, we will assume that f is continuous in (t, u, u') and of class $C^{(1)}$ with respect to u, u' for all u of interest. Equation (3.1) possesses essentially two types of solutions which are termed regular and singular. A solution $u = u(t)$ of (3.1) is called a regular solution if $f_{y'}(t, u(t), u'(t)) \not\equiv 0$, for t-values in the interval of existence, while a solution u is called singular if $f_{y'}(t, u(t), u'(t)) \equiv 0$, for such t. In this context the term "singular" is classical and refers to the fact that along singular roots, the equation $f(t, u, u') = 0$ cannot be solved for u' as a function of t and u. Consequently, it is not possible to prescribe arbitrary initial or terminal conditions for singular solutions of (3.1). On the other hand, regular solutions of (3.1) usually contain a constant of integration, and thus one boundary condition may be imposed upon these solutions, under suitable circumstances. With this as background, we are thus first of all interested in regular solutions u of (3.1) which satisfy one of the boundary conditions, $u(0) = A$ or $u(1) = B$. Let us denote regular solutions u of (3.1) which satisfy $u(0) = A$ and exist on $[0, t_L]$, $0 < t_L \leqq 1$, by the subscript "L", i.e., $u = u_L$; similarly, regular solutions u of (3.1) with $u(1) = B$ which exist on $[t_R, 1]$, $0 \leqq t_R < 1$, are denoted by u_R. Singular solutions of (3.1) will be denoted by the subscript "s", $u = u_s$.

From among this class of reduced solutions or roots (i.e., solutions of the reduced equation (3.1)), we next wish to select those which are stable in the following sense. For roots u_L, we require

$$f_{y'}[u_L(t)] = f_{y'}(t, u_L(t), u_L'(t)) \geqq 0, \qquad 0 \leqq t \leqq t_L;$$

while for roots u_R, we require

$$f_{y'}[u_R(t)] \leqq 0, \qquad t_R \leqq t \leqq 1.$$

These restrictions are simply that, in the terminology of Vasil'eva [19, Chap. 1], roots u_L are negatively stable and roots u_R are positively stable. In the case of singular roots u_s, we impose stability in the form of conditions on the partial derivatives with respect to y, $\partial_y^k f[u_s]$; recall that, by definition, $f_{y'}[u_s] \equiv 0$. Briefly we require that

$$\partial_y^{2j+1} f[u_s] \geqq 0, \qquad j = 0, 1, \cdots, N, \quad \text{for all } A \text{ and } B;$$

while if $u_s(0) < A$, $u_s(1) < B$ and $f_{y'y'} > 0$,

$$\partial_y^{2j} f[u_s] \geqq 0, \qquad j = 0, 1, \cdots, N.$$

If $u_s(0) > A$, $u_s(1) > B$, and $f_{y'y'} < 0$, stability of u_s is translated into the condition $\partial_y^{2j} f[u_s] \leqq 0$. These conditions, which will be discussed more fully later, are related to those given by Yu. P. Boglaev [1] and the author [9] in the case of the problem

$$\varepsilon^2 y'' = h(t, y), \qquad a < t < b,$$

$$y(a, \varepsilon), \qquad y(b, \varepsilon) \quad \text{prescribed.}$$

In summary, then, one of our basic assumptions is that the reduced equation (3.1) possesses regular and/or singular solutions with the stability properties outlined above. Indeed, the next three sections will be concerned with demonstrating under what conditions these reduced solutions can be used to approximate solutions of the full problem (1.1), (1.2) for all sufficiently small values of $\varepsilon > 0$.

4. Boundary layer phenomena. In this section we use the stable regular and singular reduced solutions introduced above to study solutions of the boundary value problem

$$\varepsilon y'' = f(t, y, y'), \qquad 0 < t < 1, \tag{4.1}$$

$$y(0, \varepsilon) = A, \qquad y(1, \varepsilon) = B, \tag{4.2}$$

which possess boundary layers at $t = 0$, $t = 1$ or both endpoints, for small values of $\varepsilon > 0$. We assume, for definiteness, that $f_{y'y'} < 0$ in a region defined below; at the end of this section we indicate the modifications required to handle the analogous case $f_{y'y'} > 0$.

Consider first the stable regular roots of the reduced equation, $f(t, u, u') = 0$, which we generically denote as u_L and u_R. In order for these roots to generate solutions of (4.1), (4.2) which possess boundary layers at $t = 0$ or $t = 1$, we first require these functions to exist on all of [0, 1] and to satisfy the correct stability conditions there, i.e., $f_{y'}[u_L] \geqq 0$ and $f_{y'}[u_R] \leqq 0$. More fundamental, however, is that we require the following sign restrictions on u_L and u_R; namely, $u_L(1) \geqq B$ and $u_R(0) \geqq A$. The naturalness of these conditions can be demonstrated as follows: if a solution y of (4.1), (4.2) possesses a boundary layer, for example, at $t = 1$, then in the boundary layer region adjacent to $t = 1$, the derivative of y is unbounded as a function of ε. (See [19, Chap. 2] for a precise estimate of $y'(1, \varepsilon)$.) Consequently, the "curvature" of y, y'', is represented approximately by $\varepsilon y'' \sim \frac{1}{2} f_{y'y'}(y')^2$, i.e., for small $\varepsilon > 0$, $y''(1)$ is large and negative. That is to say, inside the boundary layer region near $t = 1$, the solution is concave ($y'' < 0$). Therefore, if such a solution is to be approximated on [0, 1] by the reduced solution u_L, it must be the case that $u_L(1) \geqq B$. (If $u_L(1)=B$, then there is no boundary layer at $t=1$.) A similar argument shows why we must also require $u_R(0) \geqq A$, in case the boundary layer occurs at $t=0$. These heuristic considerations are made precise in the following theorem, stated for the case of a boundary layer at $t=0$.

THEOREM 4.1. *Assume*

1) *the reduced problem*, $f(t, u, u')=0$, $u(1)=B$, *has a solution* $u=u_R(t)$ *of class* $C^{(2)}[0, 1]$ *which satisfies* $u_R(0) \geqq A$;

2) *the function* f *is continuous in* (t, y, y') *and of class* $C^{(2)}$ *with respect to* y, y' *in* $\mathcal{R}$: $0 \leqq t \leqq 1$, $|y-u_R(t)| \leqq d$, $d \geqq |A-u_R(0)|$, $|y'|<\infty$; *also*, $f_{y'y'}=O(1)$ *in* $\mathcal{R}$;

3) *there are positive constants* k *and* p_0 *such that* $f_{y'}(t, u_R(t), u_R'(t)) \leqq -k < 0$, $0 \leqq t \leqq 1$, *and* $f_{y'y'} \leqq -p_0 < 0$ *in* $\mathcal{R}$.

Then for each $\varepsilon>0$, ε *sufficiently small, the problem* (4.1), (4.2) *has a solution* $y=y(t, \varepsilon)$. *Moreover, for* t *in* [0, 1],

$$u_R(t)-(u_R(0)-A)\exp[(c-k\varepsilon^{-1})t]-\varepsilon\bar{\gamma} \leqq y(t, \varepsilon) \leqq u_R(t)+\varepsilon\bar{\gamma}$$

where c *and* $\bar{\gamma}$ *are positive constants independent of* ε.

Proof. The theorem is proved by constructing suitable bounding functions α and β, and then applying the result of Nagumo–Jackson, Theorem 2.1. For t in [0, 1] and $\varepsilon>0$, define

$$\alpha(t, \varepsilon)=u_R(t)-(u_R(0)-A)\exp[\lambda_1 t]-\varepsilon\gamma l^{-1}(\exp[\lambda_2(t-1)]-1),$$
$$\beta(t, \varepsilon)=u_R(t)+\varepsilon\gamma l^{-1}(\exp[\lambda_2(t-1)]-1),$$

where $\lambda_1=\lambda_1(\varepsilon)<0$ (of order $O(\varepsilon^{-1})$), $\lambda_2=\lambda_2(\varepsilon)<0$ (of order $O(1)$), and $\gamma>0$ are quantities to be determined, and $|f_y(t, y, u_R'(t))| \leqq l$, $|y-u_R(t)| \leqq d$. Clearly, $\alpha \leqq \beta$, $\alpha(0, \varepsilon) \leqq A \leqq \beta(0, \varepsilon)$ and $\alpha(1, \varepsilon) \leqq B \leqq \beta(1, \varepsilon)$. To verify that α satisfies the correct differential inequality, i.e., $\varepsilon\alpha''-f(t, \alpha, \alpha') \geqq 0$, we calculate $f(t, \alpha, \alpha')$ via the Taylor expansion given at the end of § 2, with $N=1$, to obtain

$$\begin{aligned}
&\varepsilon\alpha''-f(t, \alpha, \alpha')\\
&\quad=\varepsilon u_R''-\varepsilon\lambda_1^2(u_R(0)-A)\exp[\lambda_1 t]-\varepsilon\lambda_2^2\varepsilon\gamma l^{-1}\exp[\lambda_2(t-1)]-f[u_R]\\
&\qquad\quad+f_y(t, u_R+\mathcal{O}_1(\alpha-u_R), u_R')\\
&\qquad\qquad\cdot((u_R(0)-A)\exp[\lambda_1 t]+\varepsilon\gamma l^{-1}\exp[\lambda_2(t-1)]-1))\\
&\qquad\quad+f_{y'}[u_R](\lambda_1(u_R(0)-A)\exp[\lambda_1 t]+\lambda_2\varepsilon\gamma l^{-1}\exp[\lambda_2(t-1)])\\
&\qquad\quad+f_{yy'}(t, u_R+\mathcal{O}_3(\alpha-u_R), u_R')\\
&\qquad\qquad\cdot((u_R(0)-A)\exp[\lambda_1 t]+\varepsilon\gamma l^{-1}(\exp[\lambda_2(t-1)]-1))\\
&\qquad\quad\cdot(-\lambda_1(u_R(0)-A)\exp[\lambda_1 t]-\lambda_2\varepsilon\gamma l^{-1}\exp[\lambda_2(t-1)])\\
&\qquad\quad-\frac{1}{2}f_{y'y'}(t, \alpha, u_R'+\mathcal{O}_2(\alpha'-u_R'))\\
&\qquad\qquad\cdot(\lambda_1(u_R(0)-A)\exp[\lambda_1 t]+\lambda_2\varepsilon\gamma l^{-1}\exp[\lambda_2(t-1)])^2
\end{aligned}$$

$$\begin{aligned}
&\geqq -\varepsilon M-\varepsilon\lambda_1^2(u_R(0)-A)\exp[\lambda_1 t]-\varepsilon\lambda_2^2\varepsilon\gamma l^{-1}\exp[\lambda_2(t-1)]\\
&\quad -l(u_R(0)-A)\exp[\lambda_1 t]-\varepsilon\gamma\exp[\lambda_2(t-1)]+\varepsilon\gamma\\
&\quad\quad -k\lambda_1(u_R(0)-A)\exp[\lambda_1 t]\\
&\quad -k\lambda_2\varepsilon\gamma l^{-1}\exp[\lambda_2(t-1)]+L\lambda_1(u_R(0)-A)^2\exp[2\lambda_1 t]\\
&\quad +L\lambda_2\varepsilon\gamma l^{-1}\exp[\lambda_2(t-1)]\\
&\quad\quad \cdot(u_R(0)-A)\exp[\lambda_1 t]+L\lambda\varepsilon\gamma l^{-1}(\exp[\lambda_2(t-1)]-1)\\
&\quad\quad \cdot(u_R(0)-A)\exp[\lambda_1 t]\\
&\quad +L\lambda_2(\varepsilon\gamma l^{-1})^2\exp[\lambda_2(t-1)](\exp[\lambda_2(t-1)]-1)\\
&\quad +\frac{p_0}{2}\lambda_1^2(u_R(0)-A)^2\exp[2\lambda_1 t]\\
&\quad +p_0\lambda_1\lambda_2\varepsilon\gamma l^{-1}\exp[\lambda_2(t-1)](u_R(0)-A)\exp[\lambda_1 t]\\
&\quad +\frac{p_0}{2}\lambda_2^2(\varepsilon\gamma l^{-1})^2\exp[2\lambda_2(t-1)],
\end{aligned}$$

where $|u_R''|\leqq M$ and $|f_{yy'}(t,y,u_R')|\leqq L$, $|y-u_R(t)|\leqq d$. Clearly, for $0<\varepsilon\ll 1$, since $\lambda_1=O(\varepsilon^{-1})$, $(p_0/2)\lambda_1^2\geqq -L\lambda_1$ and $p_0\lambda_1\lambda_2\geqq -L\lambda_2$. Consequently, we continue with the inequality

$$\begin{aligned}
\varepsilon\alpha''-f(t,\alpha,\alpha')\geqq &-\varepsilon M+\varepsilon\gamma\\
&-(\varepsilon\lambda_1^2+(k-\varepsilon\gamma l^{-1}L(\exp[\lambda_2(t-1)]-1))\lambda_1+l)\\
&\quad\cdot(u_R(0)-A)\exp[\lambda_1 t]\\
&-(\varepsilon\lambda_2^2+k\lambda_2+l)\varepsilon\gamma l^{-1}\exp[\lambda_2(t-1)]-K(\gamma)\varepsilon^2.
\end{aligned}$$

where

$$\begin{aligned}
&|(p_0/2)\lambda_2^2\gamma^2 l^{-2}\exp[2\lambda_2(t-1)]+L\lambda_2\gamma^2 l^{-2}\exp[\lambda_2(t-1)]\\
&\quad\cdot(\exp[\lambda_2(t-1)]-1)|\leqq K(\gamma).
\end{aligned}$$

Now choose $\lambda_2<0$ as the $O(1)$-root of $\varepsilon\lambda^2+k\lambda+l=0$, i.e., $\lambda_2=-lk^{-1}+O(\varepsilon)$. Next, pick $\gamma=M+1$; and take ε so small that (i) $k_1=k-\varepsilon\gamma l^{-1}L(\exp[\lambda_2(t-1)]-1)$ is positive and of order $O(1)$, and (ii) $\varepsilon K(\gamma)<1$. Finally, choose $\lambda_1<0$ as the $O(\varepsilon^{-1})$-root of $\varepsilon\lambda^2+k_1\lambda+l=0$, i.e., $\lambda_1=-k_1\varepsilon^{-1}+O(lk_1^{-1})+O(\varepsilon)$. Such prescriptions of the quantities satisfying these estimates can always be made for $0<\varepsilon\ll 1$. We thus have the desired inequality $\varepsilon\alpha''-f(t,\alpha,\alpha')\geqq 0$. The verification that $f(t,\beta,\beta')-\varepsilon\beta''\geqq 0$ proceeds similarly and we conclude, via Theorem 2.1, that for all sufficiently small $\varepsilon>0$, the problem (4.1), (4.2) possesses a solution $y=y(t,\varepsilon)$ with $\alpha(t,\varepsilon)\leqq y(t,\varepsilon)\leqq\beta(t,\varepsilon)$, $0\leqq t\leqq 1$.

We remark that the device of using the roots of the characteristic equation $\varepsilon\lambda^2+k\lambda+l=0$ to construct bounding functions is due to Briš [2]. It is also clear that the theorem is valid if the function f and the boundary data A, B depend in a sufficiently smooth manner on ε. Finally we note that if the function f_y satisfies the positivity assumption, $f_y(t,y,u_R')\geqq l>0$, then the estimate of Theorem 4.1 is valid with $c=0$ and k replaced by k_1. This follows most easily by defining the functions $\alpha(t,\varepsilon)=u_R(t)-(u_R(0)-A)\exp[-kt\varepsilon^{-1}]-\varepsilon\gamma l^{-1}$, $\beta(t,\varepsilon)=u_R(t)+\varepsilon\gamma l^{-1}$, and proceeding as above.

In the case when the boundary layer occurs at $t=1$, the theorem carries over verbatim with u_R replaced by $u_L\,(u_L(1)\geqq B)$ and with the inequality in assumption (3) replaced by $f_{y'}[u_L]\geqq k>0$. Of course, we still assume that $f_{y'y'}\leqq -p_0<0$ in $\mathcal{R}$. The proof of this follows by setting $\tau=1-t$ and applying Theorem 4.1 to the transformed problem.

We next consider the possibility that the stable singular solutions u_s of the reduced equation, $f(t, u, u')=0$, may be used to construct solutions of the full problem (4.1), (4.2) which possess boundary layers. As in the case of the regular reduced roots, we require the functions u_s to exist on [0, 1] and to be stable there. (The precise meaning of "stable" will be given below.) However, unlike the case just discussed, since u_s in general satisfies neither boundary condition, we anticipate boundary layer behavior at *each* end-point, and we must impose the two sign restrictions: $u_s(0)\geqq A$, $u_s(1)\geqq B$. The following theorems make these ideas precise.

THEOREM 4.2. *Assume*

1) *the reduced equation,* $f(t, u, u')=0$, *has a singular solution* $u=u_s(t)$ *of class* $C^{(2)}[0, 1]$ *with* $u_s(0)\geqq A$, $u_s(1)\geqq B$;

2) *the function f is continuous in* (t, y, y'), *of class* $C^{(2)}$ *with respect to* y', *and of class* $C^{(2q+1)}(q\geqq 0)$ *with respect to y in D*: $0\leqq t\leqq 1$, $|y-u_s(t)|\leqq d$, $d\geqq \max\{|A-u_s(0)|, |B-u_s(1)|\}$, $|y'|<\infty$; *also,* $f_{y'y'}=O(1)$ *in D, with* $f_{y'y'}\leqq -p_0<0$, *for* p_0 *a positive constant*;

3) *for t in* [0, 1], $\partial_y^j f[u_s(t)]\equiv 0$, $j=0, 1, \cdots, 2q$; $\partial_y^{2q+1} f(t, y, u_s'(t))\geqq m>0$, $|y-u_s(t)|\leqq d$, *for m a positive constant.*

Then for each $\varepsilon>0$, ε *sufficiently small, the problem* (4.1), (4.2) *has a solution* $y=y(t,\varepsilon)$. *Moreover, for t in* [0, 1],

$$u_s(t)-(u_s(0)-A)\exp[-(m\varepsilon^{-1})^{1/2}t]-(u_s(1)-B)\exp[-(m\varepsilon^{-1})^{1/2}(1-t)]-\varepsilon\bar{\gamma}$$
$$\leqq y(t,\varepsilon)\leqq u_s(t)+\varepsilon\bar{\gamma},\quad \text{if } q=0;$$

$$u_s(t)-(u_s(0)-A)(1+\sigma_1(q)\varepsilon^{-1/2}t)^{-q-1}-(u_s(1)-B)(1+\sigma_2(q)\varepsilon^{-1/2}(1-t))^{-q-1}-(\varepsilon\bar{\gamma})^{(2q+1)^{-1}}$$
$$\leqq y(t,\varepsilon)\leqq u_s(t)+(\varepsilon\bar{\gamma})^{(2q+1)^{-1}},\quad \text{if } q\geqq 1.$$

Here $\bar{\gamma}$ *is a positive constant and*

$$\sigma_1(q)=mq((q+1)(2q+1)!)^{-1/2}|A-u_s(0)|^q,$$
$$\sigma_2(q)=mq((q+1)(2q+1)!)^{-1/2}|B-u_s(1)|^q.$$

Proof. The theorem is proved by again constructing suitable bounding functions α and β. Define for t in [0, 1] and $\varepsilon>0$,

$$\alpha(t,\varepsilon)=u_s(t)-(u_s(0)-A)\exp[-(m\varepsilon^{-1})^{1/2}t]$$
$$-(u_s(1)-B)\exp[-(m\varepsilon^{-1})^{1/2}(1-t)]-\varepsilon\gamma m^{-1},\quad \text{if } q=0;$$
$$\alpha(t,\varepsilon)=u_s(t)-(u_s(0)-A)(1+\sigma_1(q)\varepsilon^{-1/2}t)^{-q-1}$$
$$-(u_s(1)-B)(1+\sigma_2(q)\varepsilon^{-1/2}(1-t))^{-q-1}-(\varepsilon\gamma m^{-1})^{(2q+1)^{-1}},\quad \text{if } q\geqq 1;$$
$$\beta(t,\varepsilon)=u_s(t)+(\varepsilon\gamma m^{-1})^{(2q+1)^{-1}},\quad \text{if } q\geqq 0.$$

The constant γ is to be chosen sufficiently large and positive. Then to show that these functions satisfy the required inequalities, we proceed as in the proof of Theorem 4.1 with the exception that in the Taylor expansion of f along α and β, we take $N = 2q+1$. The details are straightforward and are omitted; we note, however, that in this case $f_{y'}[u_s] \equiv 0$, by definition.

Assumption 3) in Theorem 4.2 provides us with a definition of stability of the singular reduced solution u_s. In the following two theorems we impose slightly weaker stability restrictions on u_s and achieve essentially the same result as in Theorem 4.2.

THEOREM 4.3. *Assume*

1) *the reduced equation, $f(t, u, u') = 0$, has a singular solution $u = u_s(t)$ of class $C^{(2)}[0, 1]$, with $u_s(0) \geqq A$, $u_s(1) \geqq B$ and $u_s'' \leqq 0$;*

2) *the function f is continuous in (t, y, y'), of class $C^{(2)}$ with respect to y' and of class $C^{(n)} (n \geqq 2)$ with respect to y in $\mathscr{D}: 0 \leqq t \leqq 1$, $|y - u_s(t)| \leqq d$, $d \geqq \max\{|A - u_s(0)|, |B - u_s(1)|\}$, $|y'| < \infty$; also, $f_{y'y'} = O(1)$ in $\mathscr{D}$ and $f_{y'y'} \leqq -p_0 < 0$, for p_0 a positive constant;*

3) *for t in $[0, 1]$,*

$$\partial_y^{2j+1} f[u_s(t)] \geqq 0, \qquad j = 0, 1, \cdots, l-1;$$
$$\partial_y^{2j} f[u_s(t)] \leqq 0, \qquad j = 0, 1, \cdots, l-1;$$

and

$$\partial_y^{2l} f(t, y, u_s'(t)) \leqq -m < 0, \quad |y - u_s(t)| \leqq d, \quad 2l \leqq n,$$

for m a positive constant.

Then for each $\varepsilon > 0$, ε sufficiently small, the problem (4.1), (4.2) *has a solution $y = y(t, \varepsilon)$. Moreover,*

$$u_s(t) - (u_s(0) - A)(1 + \sigma_1(l)\varepsilon^{-1/2} t)^{-2(2l-1)^{-1}} - (u_s(1) - B) \cdot (1 + \sigma_2(l)\varepsilon^{-1/2}(1-t))^{-2(2l-1)^{-1}}$$
$$- \varepsilon^{2(2l+1)^{-1}} \bar{\gamma} \leqq y(t, \varepsilon) \leqq u_s(t), \qquad 0 \leqq t \leqq 1,$$

where $\bar{\gamma} > 0$ is a constant and

$$\sigma_1(l) = m^{1/2}(2l-1)(2(2l+1)!)^{-1/2} |A - u_s(0)|^{(2l-1)/2},$$
$$\sigma_2(l) = m^{1/2}(2l-1)(2(2l+1)!)^{-1/2} |B - u_s(1)|^{(2l-1)/2}.$$

This theorem is proved in the same manner as Theorem 4.1 (with $N = 2l$). We note that the conclusion of Theorem 4.3 is valid if assumption 3) is replaced by 3'):

$$\partial_y^{2j+1} f[u_s] \geqq 0, \qquad j = 0, 1, \cdots, l-1;$$
$$\partial_y^{2j} f[u_s] \leqq 0, \qquad j = 0, 1, \cdots, l-1;$$

and

$$\partial_y^{2l+1} f(t, y, u_s') \geqq m > 0,$$

for m a positive constant and $|y - u_s(t)| \leqq d$.

In the two previous theorems the strict positivity or negativity of the respective partial derivatives was assumed to hold along the path (t, y, u_s') for

$|y-u_s(t)| \leqq d$. In the next theorem we require this nonvanishing to occur only along the reduced path (t, u_s, u_s').

THEOREM 4.4. *Assume* 1) *and* 2) *as in Theorem* 4.3. *Assume also* 3_1): *for* t *in* $[0, 1]$,

$$\partial_y^{2j+1} f[u_s(t)] \geqq 0, \qquad j = 0, 1, \cdots, l-1;$$
$$\partial_y^{2j} f[u_s(t)] \leqq 0, \qquad j = 0, 1, \cdots, l-1;$$
$$\partial_y^{2l} f[u_s(t)] \leqq -m < 0, \quad \textit{for } m \textit{ a positive constant},$$

and

$$\partial_y^{2l+1} f(t, y, u_s'(t)) \geqq 0, \qquad |y-u_s(t)| \leqq d.$$

Then the conclusion of Theorem 4.3 *is valid.*

We note that the conclusion of this theorem is valid if assumption 3_1) is replaced by $3_1'$):

$$\partial_y^{2j+1} f[u_s] \geqq 0, \quad \partial_y^{2j} f[u_s] \leqq 0, \quad \partial_y^{2l+1} f[u_s] \geqq m > 0,$$

for m a positive constant, *and*

$$\partial_y^{2l+2} f(t, y, u_s') \leqq 0, \qquad |y-u_s(t)| \leqq d.$$

In the case $l = 0$, the boundary layer terms are, of course, exponential functions of t and $\varepsilon^{1/2}$.

Before passing to a brief discussion of the case in which $f_{y'y'} > 0$, we consider a situation which, in a sense, is intermediate between the two cases discussed in the theorems above. Namely, it may happen that the regular roots u_L and u_R do not satisfy the strict inequalities $f_{y'}[u_L] \geqq k > 0$, $f_{y'}[u_R] \leqq -k < 0$, respectively; instead, only the inequalities $f_{y'}[u_L] \geqq 0$, $f_{y'}[u_R] \leqq 0$ may hold. The stability of such roots is then, as in the case of singular roots, determined by the signs of the partial derivatives, $\partial_y^j f[u_L]$ or $\partial_y^j f[u_R]$, for some range of $j \geqq 1$. Indeed, Theorems 4.2, 4.3, and 4.4 (and the remarks following them) apply directly to the case of regular roots u_L, u_R which satisfy $f_{y'}[u_L] \geqq 0$, $f_{y'}[u_R] \leqq 0$, respectively. Note that there is now only *one* boundary layer (at $t = 1$, if $u = u_L$; at $t = 0$, if $u = u_R$). Finally it may happen that for the regular roots u_L, u_R none of the stability assumptions in Theorems 4.1–4.4 applies; an example is given in § 8. In this case, it does not seem possible to make any general statements, and one is forced to treat such problems on an individual basis.

We conclude this section with some remarks on the case when $f_{y'y'}$ is strictly positive in the regions $\mathscr{R}$ or $\mathscr{D}$ introduced above. Quite obviously, in order that the reduced solutions $u = u_L$, u_R or u_s generate solutions of the full problem (4.1), (4.2) which possess boundary layers at the endpoints, we must have $u(0) \leqq A$ and $u(1) \leqq B$. Put geometrically, solutions of (4.1), (4.2) which possess boundary layers are, for small $\varepsilon > 0$, convex in the boundary layer regions, and the reduced solutions must reflect this convexity near the appropriate endpoint. Clearly, Theorem 4.1 is valid in this case, $f_{y'y'} \geqq p_0 > 0$, provided $u_L(1) \leqq B$ or $u_R(0) \leqq A$.

Simply define, for example in the case $u = u_R$,

$$\alpha(t, \varepsilon) = u_R(t) - \varepsilon\gamma l^{-1}(\exp[\lambda_2(t-1)] - 1),$$

$$\beta(t, \varepsilon) = u_R(t) - (u_R(0) - A)\exp[\lambda_1 t] + \varepsilon\gamma l^{-1}(\exp[\lambda_2(t-1)] - 1)$$

and proceed as in the proof of Theorem 4.1. Similarly the results of Theorems 4.2, 4.3, and 4.4 are valid if $f_{y'y'} \geqq p_0 > 0$, provided $u_s(0) \leqq A$ and $u_s(1) \leqq B$. However, in the statement of these theorems we must require that partial derivatives of even order be *nonnegative* or *positive*, and that $u_s'' \geqq 0$. These analogous results are proved most efficiently by making the change of dependent variable $y \to -y$, and applying Theorems 4.2, 4.3 and 4.4 to the transformed problem.

In the next two sections we investigate the possibility that solutions of (4.1), (4.2) possess interior nonuniformities associated with exchanges of stability between reduced solutions.

5. Haber–Levinson crossings. In this section we present a fundamental result of Haber and Levinson [6] and examine some of its implications for our study of the boundary value problem

$$\varepsilon y'' = f(t, y, y'), \qquad 0 < t < 1, \tag{5.1}$$

$$y(0, \varepsilon) = A, \qquad y(1, \varepsilon) = B. \tag{5.2}$$

The assumptions of Haber and Levinson involve the stability and algebraic properties of solutions of the corresponding reduced equation

$$0 = f(t, u, u'). \tag{5.3}$$

Namely they assumed that (5.3) has two solutions $u = u_L(t)$ and $u = u_R(t)$, which exist on $[0, t_0]$ and $[t_0, 1]$, respectively, for some t_0 in $(0, 1)$, and which satisfy $u_L(0) = A$ and $u_R(1) = B$. In addition, these roots are stable in the sense introduced above, i.e., for t in $[0, t_0]$, $f_{y'}[u_L(t)] \geqq k > 0$, and for t in $[t_0, 1]$, $f_{y'}[u_R(t)] \leqq -k < 0$, for some positive constant k. Finally these roots are assumed to intersect at t_0, i.e., $u_L(t_0) = u_R(t_0)$, with $u_L'(t_0) \neq u_R'(t_0)$; and at the crossing point t_0,

$$f(t_0, u_L(t_0), \omega)\begin{cases} > 0, & u_L'(t_0) < \omega < u_R'(t_0), \\ < 0, & u_R'(t_0) < \omega < u_L'(t_0). \end{cases}$$

Under the usual smoothness assumptions on f (i.e., f is continuous in (t, y, y') and of class $C^{(1)}$ with respect to y and y'), Haber and Levinson proved that for all sufficiently small $\varepsilon > 0$, the problem (5.1), (5.2) has a solution $y = y(t, \varepsilon)$ satisfying

$$\lim_{\varepsilon \to 0^+} y(t, \varepsilon) = \begin{cases} u_L(t), & 0 \leqq t \leqq t_0, \\ u_R(t), & t_0 \leqq t \leqq 1. \end{cases}$$

Also, they obtained an estimate for y'; namely,

$$\lim_{\varepsilon \to 0^+} y'(t, \varepsilon) = \begin{cases} u_L'(t), & 0 \leqq t \leqq t_0 - \delta, \\ u_R'(t), & t_0 + \delta \leqq t \leqq 1 \end{cases}$$

for $\delta > 0$ independent of ε. Finally this solution was shown to be unique in the sense that there is no other solution of (5.1), (5.2) which satisfies the above estimates for sufficiently small $\varepsilon > 0$.

This result says essentially that to the left of the crossing point, t_0, the solution y of (5.1), (5.2) follows the stable root $u = u_L$, while to the right of t_0, y follows $u = u_R$. In an order $O(\varepsilon)$-neighborhood of t_0, the solution undergoes a transition from u_L to u_R; however, since $u_L(t_0) = u_R(t_0)$, the nonuniformity in the transition appears first in $y'(t, \varepsilon)$. The exact order estimates for this transition, as well as complete asymptotic expansions of the solution, are given in [19, Chap. 2] and [16].

If we now specialize this result to the problem considered here, i.e., to functions f appearing in (5.1) satisfying $f_{y'y'} = O(1)$, then an interesting fact emerges. For the stable regular reduced solutions u_L and u_R considered in the previous section, we were interested in obtaining conditions under which these functions generated solutions of the full problem which possessed boundary layers at $t = 0$ or $t = 1$. The crucial assumption, in the case $f_{y'y'} < 0$, was that $u_L(1) \geqq B$ or $u_R(0) \geqq A$. However, the question naturally arises as to what happens if $u_L(1) < B$ *and* $u_R(0) < A$, i.e., if neither u_L nor u_R can generate a boundary layer solution of (5.1), (5.2). This question is answered, in most cases, by the theorem of Haber and Levinson. Namely, since $u_L(1) < B$ and $u_R(0) < A$, it is clear that u_L and u_R intersect at least once at a point in (0, 1). (Recall that by assumption $u_L(0) = A$ and $u_R(1) = B$.) At a point of intersection $t = t_0$, $u_L'(t_0) \neq u_R'(t_0)$, unless $u_L \equiv u_R$, since u_L and u_R are regular roots of $f(t, u, u') = 0$. Consequently, the assumed stability properties of u_L and u_R allow us to apply the Haber–Levinson result provided the crossing condition,

$$f(t_0, u_L(t_0), \omega)\begin{cases} >0, & u_L'(t_0) < \omega < u_R'(t_0), \\ <0, & u_R'(t_0) < \omega < u_L'(t_0), \end{cases}$$

is satisfied. If u_L and u_R intersect at several points in (0, 1), then this crossing condition serves to determine at which intersection point the solution of the full problem (for small $\varepsilon > 0$) actually transfers from one root to the other.

Up to this point we have assumed that the regular roots, u_L and u_R, existed and were stable on all of [0, 1]. However, it frequently happens that regular solutions of the equation, $f(t, u, u') = 0$, which satisfy an initial or terminal condition, exist and are stable only on $[0, t_L]$ or $[t_R, 1]$, $0 < t_L, t_R < 1$. Obviously such solutions cannot generate solutions of the full problem which possess boundary layers, since these roots either do not exist or are unstable in a neighborhood of the end-point of interest. It may nevertheless be possible to apply the Haber–Levinson theorem in such cases. To be specific, suppose that u_L exists and is stable on $[0, t_L]$, $0 < t_L < 1$, and that u_R exists and is stable on $[t_R, 1]$, $0 < t_R < 1$. Then if $t_L \geqq t_R$ and if u_L intersects u_R at a point in (t_R, t_L) (or at t_L, if $t_L = t_R$), it is only necessary to verify that the crossing condition is satisfied in order to apply the theorem of Haber and Levinson.

This discussion leads naturally to the question of what happens in the case that $t_L < t_R$, i.e., the domains of existence and/or stability of the regular roots u_L and u_R do *not* overlap, and there is no possibility of a Haber–Levinson crossing. The resolution of this problem, at least for some cases of interest, is the subject of the next section.

6. Crossings between regular and singular reduced solutions. In this section we consider the case in which stable regular reduced solutions intersect stable singular reduced solutions at one or more points in $(0, 1)$. Under appropriate assumptions, we will see that there exist solutions of the full problem which (for small $\varepsilon > 0$) follow the solution path formed from the intersection of the regular and singular roots. In addition, it is possible to have boundary layer solutions branching off of the singular roots if suitable sign restrictions are satisfied. A basic fact which must be noted here is that if $u(t_1) = u_s(t_1)$, t_1 in $(0, 1)$, where u_s is a singular solution of $f(t, u, u') = 0$ and u is any other solution (regular or singular) of this equation, then $u'(t_1) = u_s'(t_1)$. The precise result is contained in the following lemma.

LEMMA 6.1. *Let $u_s = u_s(t)$ be a singular solution of $f(t, u, u') = 0$, and let $u = u(t)$ be any other solution of this equation. Assume that for t_1 in $(0, 1)$, $u(t_1) = u_s(t_1)$ and $f_{y'y'}[t_1] \neq 0$, where $[t_1] = (t_1, u_s(t_1) + \mathcal{O}(u(t_1) - u_s(t_1)), u_s'(t_1) + \mathcal{O}(u'(t_1) - u_s'(t_1)))(0 < \mathcal{O} < 1)$, then $u'(t_1) = u_s'(t_1)$.*

Proof. The lemma follows trivially from the fact that, by definition, $f_{y'}[u_s] \equiv 0$. Indeed,

$$\begin{aligned} 0 &= f[u(t_1)] - f[u_s(t_1)] \\ &= f_y[u_s(t_1)](u(t_1) - u_s(t_1)) + f_{y'}[u_s(t_1)](u'(t_1) - u_s'(t_1)) + \tfrac{1}{2} f_{yy}[t_1](u(t_1) - u_s(t_1))^2 \\ &\quad + f_{yy'}[t_1](u(t_1) - u_s(t_1))(u'(t_1) - u_s'(t_1)) + \tfrac{1}{2} f_{y'y'}[t_1](u'(t_1) - u_s'(t_1))^2, \end{aligned}$$

i.e., $u'(t_1) = u_s'(t_1)$.

The lemma asserts that crossings between singular solutions and regular solutions, unlike Haber–Levinson crossings between regular solutions, are smooth up to first derivatives, provided $f_{y'y'}[t_1]$ is not zero at the crossing point, t_1. Such smooth joins facilitate the construction of appropriate bounding functions α and β, as will be seen below.

We turn now to a consideration of the question raised at the end of the last section; namely, what types of solutions of the problem

$$\varepsilon y'' = f(t, y, y'), \qquad 0 < t < 1, \tag{6.1}$$

$$y(0, \varepsilon) = A, \qquad y(1, \varepsilon) = B, \tag{6.2}$$

can exist when the domains of existence and/or stability of the regular solutions u_L, u_R do not overlap in $[0, 1]$. To be specific, suppose that u_L exists on $[0, t_L]$, $0 < t_L < 1$, with

$$f_{y'}[u_L(t)] \begin{cases} > 0, & 0 \leqq t < t_1, \\ = 0, & t = t_1, \\ \leqq 0, & t_1 < t \leqq t_L; \end{cases} \tag{6.3}$$

while u_R exists on $[t_R, 1]$, $0 < t_R < 1$, with

$$f_{y'}[u_R(t)] \begin{cases} < 0, & t_2 < t \leqq 1, \\ = 0, & t = t_2, \\ \geqq 0, & t_R \leqq t < t_2, \end{cases} \tag{6.4}$$

and that $t_L < t_R$. Then it may happen that there is a stable singular solution $u = u_s$ such that $u_s(t_1) = u_L(t_1)$ and $u_s(t_2) = u_R(t_2)$. If this is the case, we anticipate that for sufficiently small $\varepsilon > 0$, there exists a solution $y = y(t, \varepsilon)$ of (6.1), (6.2) such that

$$\lim_{\varepsilon\to 0^+} y(t,\varepsilon) = \begin{cases} u_L(t), & 0 \leqq t \leqq t_1, \\ u_s(t), & t_1 \leqq t \leqq t_2, \\ u_R(t), & t_2 \leqq t \leqq 1. \end{cases}$$

A verification that this does occur for several cases of interest is the content of the following theorems.

THEOREM 6.1. *Assume*

1) *there are reduced solutions u_L, u_R of class $C^{(2)}[0, 1]$ as described above satisfying* (6.3), (6.4), *respectively, with u_L'', $u_R'' \geqq 0$; there is a singular reduced solution u_s of class $C^{(2)}[0, 1]$ with $u_s'' \geqq 0$ such that $u_L(t_1) = u_s(t_1)$ and $u_R(t_2) = u_s(t_2)$, $t_1 \leqq t_L < t_R \leqq t_2$;*

2) *the function f is continuous in (t, y, y'), of class $C^{(2)}$ with respect to y' and of class $C^{(n)} (n \geqq 1)$ with respect to y in $\mathcal{D}$: $0 \leqq t \leqq 1$, $|y - u_L(t)| \leqq d$, $0 \leqq t \leqq t_1$, $|y - u_s(t)| \leqq d$, $t_1 \leqq t \leqq t_2$, $|y - u_R(t)| \leqq d$, $t_2 \leqq t \leqq 1$, $d > 0$ and possibly small, $|y'| < \infty$; $f_{y'y'} = O(1)$ in $\mathcal{D}$, $f_{y'y'} \geqq p_0 > 0$, for p_0 a positive constant; and if $n \geqq 2$, $(p_0/2)(-\lambda) \geqq L(\exp[-\lambda t_1] - 1)$, $(p_0/2)(-\lambda) \geqq L(\exp[-\lambda t_2] - 1)$, where $|f_{yy'}| \leqq L$ in $\mathcal{D}$ and $\lambda < 0$ is specified in the proof below;*

3) *for t in $[t_1 - \delta, t_2 + \delta]$, and*

$$u(t) = \begin{cases} u_L(t), & t_1 - \delta \leqq t \leqq t_1 \\ u_s(t), & t_1 \leqq t \leqq t_2, \\ u_R(t), & t_2 \leqq t \leqq t_2 + \delta, \end{cases}$$

$\partial_y^j f[u(t)] \geqq 0$, $j = 1, \cdots, n-1$, for $\delta > 0$ a small constant; moreover, $\partial_y^n f(t, y, u'(t)) \geqq m > 0$, $|y - u(t)| \leqq d$, for a positive constant m.

Then for each $\varepsilon > 0$, ε sufficiently small, the problem (6.1), (6.2) *has a solution $y = y(t, \varepsilon)$ with*

$$\begin{aligned} u_L(t) \leqq y(t,\varepsilon) \leqq u_L(t) + \varepsilon^{n^{-1}}\bar{\gamma}, &\qquad 0 \leqq t \leqq t_1, \\ u_s(t) \leqq y(t,\varepsilon) \leqq u_s(t) + \varepsilon^{n^{-1}}\bar{\gamma}, &\qquad t_1 \leqq t \leqq t_2, \\ u_R(t) \leqq y(t,\varepsilon) \leqq u_R(t) + \varepsilon^{n^{-1}}\bar{\gamma}, &\qquad t_2 \leqq t \leqq 1, \end{aligned}$$

where $\bar{\gamma} > 0$ is a constant independent of ε.

Proof. The theorem is proved by using the Habets–Laloy extension of Nagumo's theorem, Theorem 2.2. We first make some preliminary observations. Since $u_L(t_1) = u_s(t_1)$ and hence, $u_L'(t_1) = u_s'(t_1)$ by Lemma 6.1, the function $u(t)$ is continuously differentiable on $[t_1 - \delta, t_2 + \delta]$. By assumption 1), there is a constant $k_1 > 0$ such that $f_{y'}[u_L(t)] \geqq k_1 > 0$, for t in $[0, t_1 - \delta]$; similarly, there is a constant $k_2 > 0$ such that $f_{y'}[u_R(t)] \leqq -k_2 < 0$, for t in $[t_2 + \delta, 1]$. Set $k = \min\{k_1, k_2\}$. Finally let $\lambda < 0$ be the $O(1)$-root of $\varepsilon\lambda^2 + k\lambda + l = 0$, where $|f_y(t, y, u_{L,R}'(t))| \leqq l$, $|y - u_{L,R}(t)| \leqq d$, i.e., $\lambda = -lk^{-1} + O(\varepsilon)$. We are now able to define the following functions α and β, for $\varepsilon > 0$:

$$\alpha(t,\varepsilon)=\begin{cases} u_L(t), & 0\leqq t\leqq t_1,\\ u_s(t), & t_1\leqq t\leqq t_2,\\ u_R(t), & t_2\leqq t\leqq 1,\end{cases}$$

$$\beta(t,\varepsilon)=\begin{cases} u_L(t)+\varepsilon^{n^{-1}}\gamma_1 l^{-1}(\exp[-\lambda t]-1), & 0\leqq t\leqq t_1,\\ u_s(t)+(\varepsilon\gamma_2 m^{-1})^{n^{-1}}, & t_1\leqq t\leqq t_2,\\ u_R(t)+\varepsilon^{n^{-1}}\gamma_3 l^{-1}(\exp[\lambda(t-1)]-1), & t_2\leqq t\leqq 1.\end{cases}$$

The constants $\gamma_i>0$ are first chosen so that β is well defined, i.e., $\gamma_1 l^{-1}(\exp[-\lambda t_1]-1)=\gamma_3 l^{-1}(\exp[\lambda(t_2-1)]-1)=(\gamma_2 m^{-1})^{n^{-1}}$. Next we observe that $D_l\alpha(t_i)=D_r\alpha(t_i)$, $i=1,2$, since $u'_L(t_1)=u'_s(t_1)$, $u'_R(t_2)=u'_s(t_2)$; while $D_l\beta(t_i)\geqq D_r\beta(t_i)$, $i=1,2$, since $\lambda<0$. It is trivial to show that $\varepsilon\alpha''-f[\alpha]\geqq 0$ on the subintervals $(0,t_1)$, (t_1,t_2) and $(t_2,1)$, since $\{u''_L, u''_s, u''_R\}\geqq 0$ on these respective intervals. In the case of β, we first examine the subinterval $[0,t_1]$; in particular, look at $[0,t_1-\delta]$, where $\delta>0$ is the constant introduced above. Clearly,

$$\begin{aligned} f(t,\beta,\beta')-\varepsilon\beta'' &= f[u_L]+f_y(t,y,u'_L)\varepsilon^{n^{-1}}\gamma_1 l^{-1}(\exp[-\lambda t]-1)\\ &\quad +f_{y'}[u_L](-\lambda\varepsilon^{n^{-1}}\gamma_1 l^{-1}\exp[-\lambda t])\\ &\quad -f_{yy'}[\,\cdot\,]\varepsilon^{n^{-1}}\gamma_1 l^{-1}(\exp[-\lambda t]-1)\lambda\varepsilon^{n^{-1}}\gamma_1 l^{-1}\exp[-\lambda t]\\ &\quad +\tfrac{1}{2}f_{y'y'}[\,\cdot\,](\lambda\varepsilon^{n^{-1}}\gamma_1 l^{-1}\exp[-\lambda t])^2-\varepsilon u''_L\\ &\quad -\varepsilon\lambda^2\varepsilon^{n^{-1}}\gamma_1 l^{-1}\exp[-\lambda t]\\ &\geqq -\varepsilon^{n^{-1}}\gamma_1\exp[-\lambda t]+\varepsilon^{n^{-1}}\gamma_1-k\lambda\varepsilon^{n^{-1}}\gamma_1 l^{-1}\exp[-\lambda t]\\ &\quad -\varepsilon M-\varepsilon\lambda^2\varepsilon^{n^{-1}}\gamma_1 l^{-1}\exp[-\lambda t]-K_1(\gamma_1)\varepsilon^{2n^{-1}}, \end{aligned}$$

where $|f_{yy'}[\,\cdot\,]\lambda l^{-2}\exp[-\lambda t](\exp[-\lambda t]-1)\gamma_1^2|\leqq K_1(\gamma_1)$ and $0\leqq u''_L\leqq M$. Since λ is a root of $\varepsilon\lambda^2+k\lambda+l=0$, we conclude that

$$f[\beta]-\varepsilon\beta''\geqq -\varepsilon M+\varepsilon^{n^{-1}}\gamma_1-K(\gamma_1)\varepsilon^{2n^{-1}}\geqq 0,$$

for $\gamma_1=M+1$ and ε sufficiently small. Next, on the subinterval $[t_1-\delta,t_1]$, we use the assumption that $\partial_y^j f[u_L]\geqq 0$, $1\leqq j\leqq n-1$, and $\partial_y^n f(t,y,u'_L)\geqq m>0$. Again, using the Taylor expansion of $f[\beta]$ we have, for t in $[t_1-\delta,t_1]$,

$$\begin{aligned} f(t,\beta,\beta')-\varepsilon\beta'' &= f[u_L]+\sum_{j=1}^{n-1}\frac{1}{j!}\partial_y^j f[u_L](\varepsilon^{n^{-1}}\gamma_1 l^{-1}(\exp[-\lambda t]-1))^j\\ &\quad +\frac{1}{n!}\partial_y^n f(t,y,u'_L)(\varepsilon^{n^{-1}}\gamma_1 l^{-1}(\exp[-\lambda t]-1))^n+f_{y'}[u_L]\\ &\quad \cdot(-\lambda\varepsilon^{n^{-1}}\gamma_1 l^{-1}\exp[-\lambda t])\\ &\quad +f_{yy'}[\,\cdot\,](-\lambda)\varepsilon^{n^{-1}}\gamma_1 l^{-1}\exp[-\lambda t]\varepsilon^{n^{-1}}\gamma_1 l^{-1}(\exp[-\lambda t]-1)\\ &\quad +\frac{1}{2}f_{y'y'}[\,\cdot\,](\lambda\varepsilon^{n^{-1}}\gamma_1 l^{-1}\exp[-\lambda t])^2-\varepsilon u''_L\\ &\quad -\varepsilon\lambda^2\varepsilon^{n^{-1}}\gamma_1 l^{-1}\exp[-\lambda t] \end{aligned}$$

$$\geqq \varepsilon \frac{m}{n!}\gamma_1^n l^{-n}(\exp[-\lambda(t_1-\delta)]-1)^n - \varepsilon M - \varepsilon\lambda^2\varepsilon^{n-1}\gamma_1 l^{-1}\exp[-\lambda t]$$

$$\geqq 0, \quad \text{if } \gamma_1 = \left\{\frac{n!}{m}l^n(\exp[-\lambda(t_1-\delta)]-1)^{-n}(M+1)\right\}^{n-1}$$

and ε is sufficiently small. Thus, $f[\beta]-\varepsilon\beta''\geqq 0$, for t in $[0, t_1]$. For $[t_1, t_2]$ the verification follows in a straightforward manner:

$$f(t,\beta,\beta')-\varepsilon\beta'' = f[u_s] + \sum_{j=1}^{n-1}\frac{1}{j!}\partial_y^j f[u_s]((\varepsilon\gamma_2 m^{-1})^{n^{-1}})^j + \frac{1}{n!}\partial_y^n f(t, y, u_s')$$

$$\cdot((\varepsilon\gamma_2 m^{-1})^{n^{-1}})^n - \varepsilon u_s''$$

$$\geqq \frac{\gamma_2}{n!}\varepsilon - \varepsilon M, \quad \text{where } 0 \leqq u_s'' \leqq M,$$

$$\geqq 0, \quad \text{for } \gamma_2 \geqq n!M.$$

The verification that β satisfies the same inequality on $[t_2, 1]$ is similar to that for $[0, t_1]$ and is omitted. We note that there is a slight technicality regarding the specification of the constants γ_i in the definition of β. They were first chosen so that β was well defined and then they were chosen to be larger than certain constants. To be precise, we required that $\gamma_1 \geqq \max\{M+1, ((n!/m)l^n(\exp[-\lambda(t_1-\delta)]-1)^{-n}(M+1))^{n-1}\}$ and $\gamma_2 \geqq n!M$. However, γ_1 and γ_2 may also be chosen so that $\gamma_1 l^{-1}(\exp[-\lambda t]-1) = (\gamma_2 m^{-1})^{n^{-1}}$; these are the γ_1, γ_2, to be used in the definition of β. A similar statement applies to γ_3.

The case in which $\{u_L'', u_s'', u_R''\} \leqq 0$ can be treated in the same way by requiring that $f_{y'y'} \leqq -p_0 < 0$ in $\mathscr{D}$ and by making the appropriate sign changes in assumption 3). That is, partial derivatives of even order must be nonpositive, and if $n \geqq 2$, $f_{y'y'}$ must be sufficiently negative to balance the term arising from $f_{yy'}$ in the expansion of f.

In the event that neither $\{u_L'', u_s'', u_R''\} \geqq 0$ nor $\{u_L'', u_s'', u_R''\} \leqq 0$, a result analogous to Theorem 6.1 appears to be provable only under the assumption that $f_y(t, y, u_s') \geqq m > 0$, $|y-u_s(t)| \leqq d$. It also seems that for the examples we have considered (some of which are given in § 8), this is the only assumption regarding f_y which guarantees that u_L, u_s, u_R-crossings do in fact occur. The precise result is the following theorem.

THEOREM 6.2. *Assume* 1) *and* 2) *as in Theorem* 6.1 *with the exception that* u_L'', u_s'' *and* u_R'' *are not required to be nonnegative; also suppose* $n = 1$ *in* 2). *Assume finally that* $f_y(t, y, u') \geqq m > 0$, *for a positive constant* m, *with* (t, y) *in* $\mathscr{D}$ *and* $u' = u_L', u_s', u_R'$, *as* t *ranges over* $[0, t_1]$, $[t_1, t_2]$, $[t_2, 1]$, *respectively.*

Then for each $\varepsilon > 0$, ε *sufficiently small, the problem* (6.1), (6.2) *has a solution* $y = y(t, \varepsilon)$ *with*

$$u_L(t)-\varepsilon\bar\gamma \leqq y(t,\varepsilon) \leqq u_L(t)+\varepsilon\bar\gamma, \qquad 0 \leqq t \leqq t_1,$$

$$u_s(t)-\varepsilon\bar\gamma \leqq y(t,\varepsilon) \leqq u_s(t)+\varepsilon\bar\gamma, \qquad t_1 \leqq t \leqq t_2,$$

$$u_R(t)-\varepsilon\bar\gamma \leqq y(t,\varepsilon) \leqq u_R(t)+\varepsilon\bar\gamma, \qquad t_2 \leqq t \leqq 1,$$

where $\bar\gamma$ *is a positive constant.*

Proof. The theorem is proved by defining for $\varepsilon > 0$ the functions

$$\alpha(t,\varepsilon)=\begin{cases} u_L(t)-\varepsilon\gamma m^{-1}, & 0\leqq t\leqq t_1,\\ u_s(t)-\varepsilon\gamma m^{-1}, & t_1\leqq t\leqq t_2,\\ u_R(t)-\varepsilon\gamma m^{-1}, & t_2\leqq t\leqq 1,\end{cases}$$

$$\beta(t,\varepsilon)=\begin{cases} u_L(t)+\varepsilon\gamma m^{-1}, & 0\leqq t\leqq t_1,\\ u_s(t)+\varepsilon\gamma m^{-1}, & t_1\leqq t\leqq t_2,\\ u_R(t)+\varepsilon\gamma m^{-1}, & t_2\leqq t\leqq 1,\end{cases}$$

and verifying that the hypotheses of Theorem 2.2 are satisfied. We omit the details.

Although we will not prove it here, it is true that since $u_L'(t_1)=u_s'(t_1)$ and $u_R'(t_2)=u_s'(t_2)$ whenever $u_L(t_1)=u_s(t_1)$ and $u_R(t_2)=u_s(t_2)$, the derivative $y'(t,\varepsilon)$ converges uniformly to u_L', u_s' and u_R' on the respective intervals $[0, t_1]$, $[t_1, t_2]$ and $[t_2, 1]$.

The next type of crossing between stable regular and singular reduced solutions which we examine involves the case in which a regular solution, u_L or u_R, crosses a singular solution u_s, and u_s generates a solution of (6.1), (6.2) possessing a boundary layer branching from u_s. A typical example of this is the following. Assume $f_{y'y'} < 0$ and that for the left-hand boundary value A, a regular solution u_L either does not exist or is unstable to the right of $t = 0$; suppose also that a stable root u_R exists on a subinterval $[t_R, 1]$, $t_R > 0$. Then if there exists a stable singular root u_s which crosses u_R at t_2, $t_R < t_2 < 1$, and if $u_s(0)\geqq A$ and u_s has compatible stability properties (cf. § 4), we anticipate that for sufficiently small $\varepsilon > 0$, a solution $y = y(t, \varepsilon)$ of (6.1), (6.2) exists and satisfies

$$\lim_{\varepsilon\to 0^+} y(t,\varepsilon)=\begin{cases} u_s(t), & 0<t\leqq t_2,\\ u_R(t), & t_2\leqq t\leqq 1.\end{cases}$$

The following theorems examine precisely under what conditions such behavior does in fact occur. We will assume, for definiteness, that $f_{y'y'} < 0$ in the region defined below; also, we examine crossings between u_R and u_s. Extensions to the other cases of interest will be indicated below.

THEOREM 6.3. *Assume*

1) *the reduced equation $f(t, u, u') = 0$ has solutions $u = u_s(t)$ of class $C^{(2)}[0, t_2]$ and $u = u_R(t)$ of class $C^{(2)}[t_2, 1]$, $0 < t_2 < 1$, with u_s'', $u_R''\leqq 0$, and such that $u_s(t_2) = u_R(t_2)$, $u_R(1) = B$ and $u_s(0)\geqq A$;*

2) *the function f is continuous in (t, y, y'), of class $C^{(2)}$ with respect to y', and of class $C^{(2q+1)}$ $(q\geqq 0)$ with respect to y in $\mathscr{D}$: $0\leqq t\leqq 1$, $|y-u_s(t)|\leqq d_1$, $0\leqq t\leqq t_2$, $|y-u_R(t)|\leqq d_2$, $t_2\leqq t\leqq 1$, $|y'|<\infty$, with $d_1\geqq u_s(0)-A$; also, $f_{y'y'}=O(1)$ in $\mathscr{D}$, $f_{y'y'}\leqq -p_0<0$, for p_0 a positive constant, and if $q\geqq 1$, $(p_0/2)(-\lambda)\geqq L(\exp[\lambda(t_2-1)]-1)$, where $|f_{yy'}|\leqq L$ in $\mathscr{D}$ and $\lambda<0$ is a quantity defined in the proof below;*

3) *for t in $[t_2, 1]$,*

$$f_{y'}[u_R(t)]\begin{cases} <0, & t_2<t\leqq 1,\\ =0, & t=t_2;\end{cases}$$

4) *for t in* $[0, t_2]$,

$$\partial_y^{2j+1} f[u_s(t)] \geqq 0, \qquad j = 0, 1, \cdots, q-1;$$

$$\partial_y^{2j} f[u_s(t)] \leqq 0, \qquad j = 0, 1, \cdots, q;$$

and

$$\partial_y^{2q+1} f(t, y, u_s'(t)) \geqq m > 0, \qquad |y - u_s(t)| \leqq d_1,$$

for a positive constant m.

The for each $\varepsilon > 0$, ε *sufficiently small, the problem* (6.1), (6.2) *has a solution* $y = y(t, \varepsilon)$. *Moreover,*

$$u_s(t) - (u_s(0) - A)\exp[-(m\varepsilon^{-1})^{1/2} t] - \varepsilon\bar{\gamma} \leqq y(t, \varepsilon) \leqq u_s(t), \qquad 0 \leqq t \leqq t_2, \quad \text{if } q = 0;$$

$$u_s(t) - (u_s(0) - A)(1 + \sigma(q)\varepsilon^{-1/2} t)^{-q^{-1}} - \varepsilon^{(2q+1)^{-1}} \bar{\gamma} \leqq y(t, \varepsilon) \leqq u_s(t), \qquad 0 \leqq t \leqq t_2, \quad \text{if } q \geqq 1;$$

$$u_R(t) - \varepsilon^{(2q+1)^{-1}} \bar{\gamma} \leqq y(t, \varepsilon) \leqq u_R(t), \qquad t_2 \leqq t \leqq 1, \quad \text{if } q \geqq 0,$$

where $\sigma(q) = \sigma_1(q)$ *of Theorem* 4.2 *and* $\bar{\gamma} > 0$ *is a constant.*

Proof. The proof follows essentially as that of Theorem 6.1 by verifying that on the subintervals $[0, t_2]$ and $[t_2, 1]$, the functions α and β defined below satisfy the required inequalities. Namely, for $q = 0$ and $\varepsilon > 0$, set

$$\alpha(t, \varepsilon) = \begin{cases} u_s(t) - (u_s(0) - A)\exp[-(m\varepsilon^{-1})^{1/2} t] \\ \qquad - t(u_s(0) - A)(m\varepsilon^{-1})^{1/2} \exp[-(m\varepsilon^{-1})^{1/2} t_2] - \varepsilon\gamma_1 m^{-1}, \quad 0 \leqq t \leqq t_2, \\ u_R(t) - (u_s(0) - A)\exp[-(m\varepsilon^{-1})^{1/2} t_2] \\ \qquad - t_2(u_s(0) - A)(m\varepsilon^{-1})^{1/2} \exp[-(m\varepsilon^{-1})^{1/2} t_2] \\ \qquad\qquad - \varepsilon\gamma_2 l^{-1}(\exp[\lambda(t-1)] - 1), \quad t_2 \leqq t \leqq 1; \end{cases}$$

while for $q \geqq 1$, set

$$\alpha(t, \varepsilon) = \begin{cases} u_s(t) - (u_s(0) - A)(1 + \sigma(q)\varepsilon^{-1/2} t)^{-q^{-1}} \\ \qquad - t(u_s(0) - A)\sigma(q)\varepsilon^{-1/2}(1 + \sigma(q)\varepsilon^{-1/2} t_2)^{-q^{-1}-1} \\ \qquad - (\varepsilon\gamma_1 m^{-1})^{(2q+1)^{-1}}, \quad 0 \leqq t \leqq t_2; \\ u_R(t) - (u_s(0) - A)(1 + \sigma(q)\varepsilon^{-1/2} t_2)^{-q^{-1}} \\ \qquad - t_2(u_s(0) - A)\sigma(q)\varepsilon^{-1/2}(1 + \sigma(q)\varepsilon^{-1/2} t_2)^{-q^{-1}-1} \\ \qquad - \varepsilon^{(2q+1)^{-1}} \gamma_2 l^{-1}(\exp[\lambda(t-1)] - 1), \quad t_2 \leqq t \leqq 1; \end{cases}$$

and for $q \geqq 0$,

$$\beta(t, \varepsilon) = \begin{cases} u_s(t), & 0 \leqq t \leqq t_2, \\ u_R(t), & t_2 \leqq t \leqq 1. \end{cases}$$

Here $\lambda < 0$ is the $O(1)$-root of $\varepsilon\lambda^2 + k\lambda + l \doteq 0$, where $|f_y(t, y, u_R'(t))| \leqq l$ and $f_{y'}[u_R(t)] \leqq -k < 0$ on $[t_2 + \delta, 1]$, $\delta > 0$. The constants $\gamma_1, \gamma_2 > 0$ are chosen to be sufficiently large and to make α well-defined, i.e.,

$$(\gamma_1 m^{-1})^{(2q+1)^{-1}} = \gamma_2 l^{-1}(\exp[\lambda(t_2-1)]-1).$$

We omit the details of the proof.

An analogous result is obviously true in the case that a stable root u_L crosses a singular solution u_s at a point t_1 in (0, 1). The assumptions are that

$$f_{y'}[u_L(t)]\begin{cases} >0, & 0 \leqq t < t_1, \\ =0, & t = t_1, \end{cases}$$

and that $u_s(1) \geqq B$. We also remark that this theorem is valid if for some integer $l \geqq 1$, $\partial_y^{2l} f \leqq -m < 0$, i.e., $\partial_y^{2j+1} f[u_s] \geqq 0$ and $\partial_y^{2j} f[u_s] \leqq 0$, $j = 1, \cdots, l-1$; and $\partial_y^{2l} f(t, y, u_s') \leqq -m < 0$, $|y - u_s(t)| \leqq d_1$, t in $[0, t_2]$. Finally in the case that u_R'' (or u_L'') and u_s'' are nonnegative *and* $f_{y'y'} > 0$ in the region $\mathscr{D}$, the result corresponding to Theorem 6.3 can also be proved.

We next consider crossings between roots u_L or u_R, and singular roots u_s in the case that neither $\{u_L'', u_R'', u_s''\} \leqq 0$ nor $\{u_L'', u_R'', u_s''\} \geqq 0$ holds. As discussed prior to the statement of Theorem 6.2, for such reduced solutions, the only reasonable assumption is that $f_y(t, y, u') \geqq m > 0$, in the appropriate domains. The precise result is the following.

THEOREM 6.4. *Assume* 1), 2) (*with* $q = 0$) *and* 3) *as in Theorem* 6.3 *with the exception that* u_s'', u_R'' *are not required to be nonpositive. Assume also that* $f_y(t, y, u'(t)) \geqq m > 0$, *for a positive constant m, where* (t, y) *belongs to* $\mathscr{D}$ *and* $u' = u_R', u_s'$ *for t in* $[t_2, 1]$, $[0, t_2]$, *respectively.*

Then for each $\varepsilon > 0$, ε *sufficiently small, the problem* (6.1), (6.2) *has a solution* $y = y(t, \varepsilon)$ *with*

$$u_s(t) - (u_s(0) - A)\exp[-(\sigma\varepsilon^{-1})^{1/2} t] - \varepsilon\bar{\gamma} \leqq y(t, \varepsilon) \leqq u_s(t) + \varepsilon\bar{\gamma}, \qquad 0 \leqq t \leqq t_2;$$

$$u_R(t) - \varepsilon\bar{\gamma} \leqq y(t, \varepsilon) \leqq u_R(t) + \varepsilon\bar{\gamma}, \qquad t_2 \leqq t \leqq 1,$$

where $0 < \sigma < m$ *and* $\bar{\gamma} > 0$ *is a constant.*

Proof. Simply define, for $\varepsilon > 0$,

$$\alpha(t, \varepsilon) = \begin{cases} u_s(t) - (u_s(0) - A)\exp[-(\sigma\varepsilon^{-1})^{1/2} t] & \\ \qquad - t(u_s(0) - A)(\sigma\varepsilon^{-1})^{1/2}\exp[-(\sigma\varepsilon^{-1})^{1/2} t_2] - \varepsilon\gamma m^{-1}, & 0 \leqq t \leqq t_2, \\ u_R(t) - t_2(u_s(0) - A)(\sigma\varepsilon^{-1})^{1/2}\exp[-(\sigma\varepsilon^{-1})^{1/2} t_2] - \varepsilon\gamma m^{-1}, & t_2 \leqq t \leqq 1; \end{cases}$$

$$\beta(t, \varepsilon) = \begin{cases} u_s(t) + \varepsilon\gamma m^{-1}, & 0 \leqq t \leqq t_2, \\ u_R(t) + \varepsilon\gamma m^{-1}, & t_2 \leqq t \leqq 1, \end{cases}$$

and verify that these functions satisfy the inequalities of Theorem 2.2.

Clearly, if $f_{y'y'} > 0$, an analogous result holds (essentially with the roles of α and β interchanged) provided $u_s(0) \leqq$ A. Similarly, the case when u_L crosses u_s can be treated by making the change of variable $\tau = 1 - t$ and applying Theorem 6.4 to the transformed problem.

With this discussion we end our consideration of the various types of asymptotic behavior which solutions of the full problem can exhibit for small values of ε. In the next section we summarize the results obtained thus far in the form of an algorithm which clarifies how the phenomena observed above arise in studying a specific problem. This algorithm can also serve as a means of studying related problems not explicitly treated here by any of the various asymptotic methods currently available.

7. A solution algorithm. In this section we present an algorithm for the asymptotic solution (as $\varepsilon \to 0^+$) of the boundary value problem we have been considering

$$\varepsilon y'' = f(t, y, y'), \qquad 0 < t < 1, \tag{7.1}$$

$$y(0, \varepsilon) = A, \qquad y(1, \varepsilon) = B, \tag{7.2}$$

where f is a smooth function of its arguments and $f_{y'y'} = O(1)$, $f_{y'y'} \neq 0$, in the domain of interest. The algorithm consists of a sequence of steps which selectively determines the various roots of the reduced equation $f(t, u, u') = 0$ having the properties which generate solutions of (7.1), (7.2) behaving in a definite manner as $\varepsilon \to 0^+$. Indeed, §§ 4–6 above contain justifications of the implications of the algorithm for a class of problems of the form (7.1), (7.2) which is especially amenable to treatment by the method of differential inequalities. We anticipate that the algorithm will also be of use in studying broader classes of such problems, for example, by formal methods like matched expansions or two-variable techniques. The conditions imposed below not surprisingly share an affinity with various types of matching relations which form the basis of these more formal techniques.

A solution algorithm.

Step 1. Find all solutions of the reduced equation

$$f(t, u, u') = 0. \tag{7.3}$$

For the regular solutions u of (7.3), i.e., those satisfying $f_{y'}[u] \not\equiv 0$, impose the boundary data (7.2) at $t = 0$ and $t = 1$. Denote the regular solutions u of (7.3) which satisfy $u(0) = A$ by u_L, i.e., $f[u_L(t)] \equiv 0$, $f_{y'}[u_L(t)] \not\equiv 0$, $0 \leqq t \leqq t_L$, $0 < t_L \leqq 1$, and $u_L(0) = A$. Similarly, denote the regular solutions u of (7.3) satisfying $u(1) = B$ by u_R, i.e., $f[u_R(t)] = 0$, $f_{y'}[u_R(t)] \not\equiv 0$, $t_R \leqq t \leqq 1$, $0 \leqq t_R < 1$, and $u_R(1) = B$. Finally denote the singular solutions of (7.3) by u_s, i.e., $f[u_s] \equiv 0$ and $f_{y'}[u_s] \equiv 0$ on the interval of existence of u_s. In general, it is not possible to prescribe any initial or terminal conditions for u_s.

Step 2. Determine the stability of the reduced solutions u_L, u_R and u_s. Namely, roots u_L are stable if $f_{y'}[u_L] \geqq 0$ in a t-interval including $t = 0$; roots u_R are stable if $f_{y'}[u_R] \leqq 0$ in a t-interval including $t = 1$. In the case of singular roots u_s, stability is determined by the sign of $f_{y'y'}$. If $f_{y'y'} > 0$ in the domain of interest, u_s is stable if $\partial_y^{2j+1} f[u_s] \geqq 0$ and $\partial_y^{2j} f[u_s] \geqq 0$, $j = 0, 1, \cdots, N$, N depending on the problem. If $f_{y'y'} < 0$, u_s is stable if $\partial_y^{2j+1} f[u_s] \geqq 0$ and $\partial_y^{2j} f[u_s] \leqq 0$, $j = 0, 1, \cdots, N$. Consider from now on only stable roots u_L, u_R and u_s.

Step 3. Determine what roots u_L, u_R or u_s generate, for small $\varepsilon > 0$, solutions $y = y(t, \varepsilon)$ of (7.1), (7.2) which possess boundary layers at the end-points, i.e., solutions such that $\lim_{\varepsilon \to 0^+} y(t, \varepsilon) = u_L(t)$, $0 \leqq t < 1$, $\lim_{\varepsilon \to 0^+} y(t, \varepsilon) = u_R(t)$, $0 < t \leqq 1$, or $\lim_{\varepsilon \to 0^+} y(t, \varepsilon) = u_s(t)$ $0 < t < 1$. (Of course, these roots must necessarily exist on [0, 1].) Namely, if $f_{y'y'}>0$, it is required that $u_L(1)\leqq B$, $u_R(0)\leqq A$, $u_s(0)\leqq A$ *and* $u_s(1)\leqq B$. Note that if one of these inequalities is actually an equality, then there is no boundary layer at that endpoint. Similarly, if $f_{y'y'}<0$, it is required that $u_L(1)\geqq B$, $u_R(0)\geqq A$, $u_s(0)\geqq A$ *and* $u_s(1)\geqq B$.

Step 4. Determine the Haber–Levinson crossings, i.e., determine if the Haber–Levinson theorem quoted in § 5 can be applied. This is always necessary in the case of roots u_L and u_R in Step 3 which satisfy *neither* sign condition relative to the boundary data. Namely if $f_{y'y'}>0$ and $u_L(1)>B$, $u_R(0)>A$, then there exists at least one point t_0 in (0, 1) at which $u_L(t_0)=u_R(t_0)$ and $u'_L(t_0)\neq u'_R(t_0)$. It is then only necessary to verify that the crossing condition is satisfied, i.e., that

$$f(t_0, u_L(t_0), \omega)\begin{cases} >0, & u'_L(t_0)<\omega<u'_R(t_0), \\ <0, & u'_R(t_0)<\omega<u'_L(t_0). \end{cases}$$

Similarly, if $f_{y'y'}<0$ and $u_L(1)<B$, $u_R(0)<A$, the occurrence of a Haber–Levinson crossing must be investigated.

Haber–Levinson crossings can also occur in the case that the domains of stability of u_L and u_R, say $[0, t_L]$, $[t_R, 1]$, respectively, $0<t_L$, $t_R<1$, overlap, i.e., $t_L\geqq t_R$. Then if $u_L(t_0)=u_R(t_0)$, t_0 in (t_R, t_L) (or $u_L(t_0=t_L=t_R)=u_R(t_0)$), the crossing condition must be checked.

Step 5. Determine whether crossings between regular and singular solutions occur. In particular, such crossings are likely when the intervals of stability of u_L and u_R do not overlap, i.e., $t_L<t_R$, where

$$f_{y'}[u_L(t)]\begin{cases} >0, & 0\leqq t<t_L, \\ =0, & t=t_L, \end{cases} \quad \text{and} \quad f_{y'}[u_R(t)]\begin{cases} <0, & t_R<t\leqq 1, \\ =0, & t=t_R. \end{cases}$$

Then if there are points t_1 in $(0, t_L]$ and t_2 in $[t_R, 1)$ at which $u_L(t_1)=u_s(t_1)$ and $u_R(t_2)=u_s(t_2)$, where u_s is a stable singular reduced solution, it is likely that for small $\varepsilon>0$, the full problem (7.1), (7.2) has a solution $y=y(t,\varepsilon)$ which follows the path formed by u_L, u_s and u_R, i.e.,

$$\lim_{\varepsilon\to 0^+} y(t,\varepsilon)=\begin{cases} u_L(t), & 0\leqq t\leqq t_1, \\ u_s(t), & t_1\leqq t\leqq t_2. \\ u_R(t), & t_2\leqq t\leqq 1. \end{cases}$$

Another type of crossing which can occur involves a regular root, say u_R, and a singular root u_s; in particular, suppose u_L either does not exist or is unstable to the right of $t=0$. Suppose also that u_R loses stability at t_R, $0<t_R<1$, and that $u_R(t_2)=u_s(t_2)$, $t_2\geqq t_R$. Then if the stability of u_s is compatible with the sign of $f_{y'y'}$, it is likely that there exists a solution of (7.1), (7.2) for small $\varepsilon>0$ which follows the path formed by u_s and u_R, with a boundary layer occurring at $t=0$. To be precise, if $f_{y'y'}>0$, it is required that $\partial_y^{2j+1}f[u_s]\geqq 0$, $\partial_y^{2j}f[u_s]\geqq 0$, *and* $u_s(0)\leqq A$; while if $f_{y'y'}<0$, it is required that $\partial_y^{2j+1}f[u_s]\geqq 0$, $\partial_y^{2j}f[u_s]\leqq 0$, *and* $u_s(0)\geqq A$, for appropriate ranges of $j\geqq 0$. Similarly, if u_R does not exist or is unstable to the left of $t=1$, the possibility of crossings between u_L and a stable singular solution should be investigated. Finally, in more complicated problems, crossings between u_L, u_R and more than one singular solution may occur.

This concludes our outline of the algorithm for the solution of the problem (7.1), (7.2), for small $\varepsilon > 0$, under the assumptions that $f_{y'y'} = O(1)$ and $f_{y'y'}$ is never zero. We remark that a similar algorithm for the analytical solution of singularly perturbed nth order linear and second-order quasilinear boundary value problems was given recently by O'Malley [18]. His algorithm has also been implemented to solve such problems numerically; some computational results are summarized in [5].

8. Some examples. We present now some examples which illustrate the theory of §§ 4–6 and the application of the algorithm given in § 7.

Example 8.1. The boundary value problem

$$\varepsilon y'' = 1 - (y')^2, \qquad 0 < t < 1, \tag{8.1}$$

$$y(0, \varepsilon) = A, \qquad y(1, \varepsilon) = B, \tag{8.2}$$

was cited by Haber and Levinson [6] as an illustration of their theorem quoted in §5, for values of A and B satisfying $0 < |A - B| < 1$. It was also considered by Dorr, Parter and Shampine [4, § 5] for such A and B, and also for $B - A \geqq 1$. Forgetting that this problem can be solved by quadratures for all A and B, we apply the algorithm of the previous section. We begin by solving the reduced problems

$$0 = 1 - (u')^2, \quad u(0) = A \quad \text{or} \quad u(1) = B, \tag{8.3}$$

and find that, in our terminology, $u_L(t) = A \pm t$, $u_R(t) = B \mp 1 \pm t$. Clearly (8.3) has no singular solutions. Setting $f(t, y, y') = 1 - (y')^2$, we calculate $f_{y'} = -2y'$, and easily determine that $u_L(t) = A - t$ and $u_R(t) = t + B - 1$ are the stable roots of (8.3), for *all* A and B. If we now apply the boundary layer criteria of Step 3 which in this case are that $u_L(1) = A - 1 \geqq B$ and $u_R(0) = B - 1 \geqq A$, we can conclude by Theorem 4.1 that for such A and B, the problem (8.1), (8.2) has, for $\varepsilon > 0$ sufficiently small, solutions $y = y(t, \varepsilon)$ satisfying (i) $\lim_{\varepsilon \to 0^+} y(t, \varepsilon) = A - t$, $0 \leqq t (\leqq) 1$, if $A - 1 (\geqq) B$, and (ii) $\lim_{\varepsilon \to 0^+} y(t, \varepsilon) = t + B - 1$, $0 (\leqq) t \leqq 1$, if $B - 1 (\geqq) A$. Indeed, for (i) we have the estimate

$$A - t - (A - 1 - B) \exp[-2\varepsilon^{-1}(1 - t)] \leqq y(t, \varepsilon) \leqq A - t, \qquad 0 \leqq t \leqq 1;$$

while, for (ii),

$$t + B - 1 - (B - 1 - A) \exp[-2t\varepsilon^{-1}] \leqq y(t, \varepsilon) \leqq t + B - 1, \qquad 0 \leqq t \leqq 1.$$

Next we examine (8.1), (8.2) for A, B-values leading to a Haber–Levinson crossing. In particular, suppose $u_L(1) = A - 1 < B$ and $u_R(0) = B - 1 < A$, then there exists in this case a single point t_0 in $(0, 1)$ at which $u_L(t_0) = u_R(t_0)$, i.e., $t_0 = \frac{1}{2}(1 - (B - A))$. It is also clear that the crossing condition is satisfied, i.e., $f(t_0, u_L(t_0), \omega) = 1 - \omega^2 > 0$ for $u_L'(t_0) = -1 < \omega < 1 = u_R'(t_0)$. For such A and B, which are characterized by the inequality $|A - B| < 1$, there is then a Haber–Levinson crossing at $t_0 = \frac{1}{2}(1 - (B - A))$. Thus the asymptotic behavior of solutions of (8.1), (8.2) has been determined for all choices of A and B, as may be seen from the following boundary value portrait. (See Fig. 1.) For example, in Figure 1, the notation "u_R + B.L.$(0, \varepsilon)$" in the region $B \geqq A + 1$ is a shorthand statement that for such A and B, the solution of (8.1), (8.2) is described asymptotically by

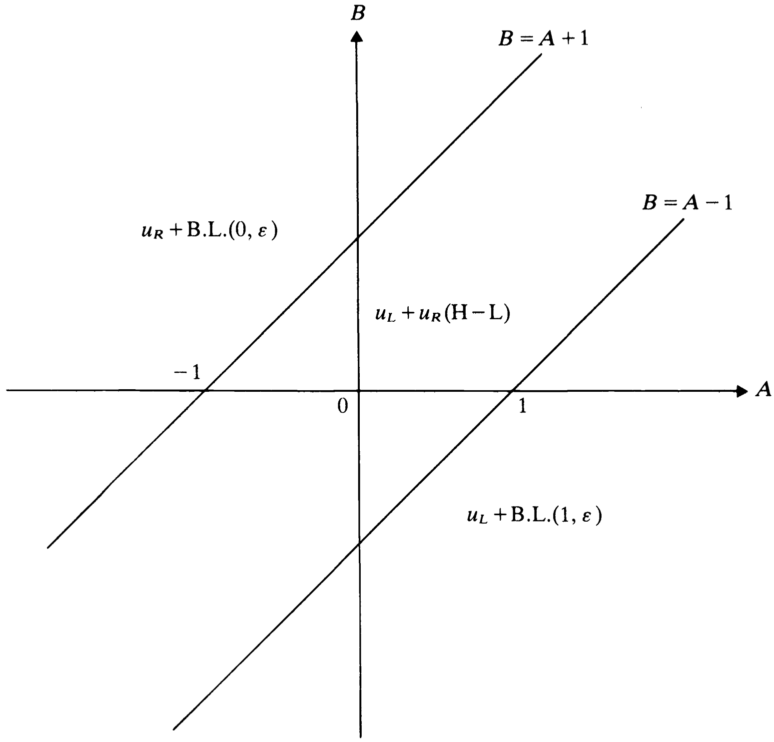

FIG. 1

the function $u_R(t) = t + B - 1$ plus a boundary layer function of ε near $t = 0$. The letters "H–L" denote that a Haber–Levinson crossing occurs for A and B in the region indicated.

Example 8.2. We consider next the problem

$$\varepsilon y'' = -yy' - (y')^2, \qquad 0 < t < 1, \tag{8.4}$$

$$y(0, \varepsilon) = A, \qquad y(1, \varepsilon) = B, \tag{8.5}$$

which was discussed by Dorr, Parter and Shampine [4, § 5] in the case that $A < B$. Our first step is naturally to determine the roots of the reduced equation

$$0 = uu' + (u')^2, \tag{8.6}$$

and we find that $u_L(t) = A$, $\bar{u}_L(t) = Ae^{-t}$, $u_R(t) = B$, $\bar{u}_R(t) = Be^{1-t}$. Again the reduced equation has no singular solutions. Setting $f(t, y, y') = -yy' - (y')^2$, we find that $f_{y'} = -y - 2y'$; consequently, $f_{y'}[u_L] = -A$, $f_{y'}[\bar{u}_L] = Ae^{-t}$, $f_{y'}[u_R] = -B$, and $f_{y'}[\bar{u}_R] = Be^{1-t}$. Thus the stability of these reduced solutions, unlike those of (8.3), depends critically on the signs of A and B.

Case 1. $A, B > 0$. Clearly for such A and B, $\bar{u}_L(t) = Ae^{-t}$ and $u_R(t) = B$ are the stable roots of (8.6). We first check the sign restrictions necessary for boundary layer behavior, i.e., $\bar{u}_L(1) = Ae^{-1} \geqq B$ and $u_R(0) = B \geqq A$. For these A and B, we conclude by Theorem 4.1 that for sufficiently small $\varepsilon > 0$, there exists a solution

$y=y(t,\varepsilon)$ of (8.4), (8.5) such that $B-(B-A)\exp[-Bt\varepsilon^{-1}]\leqq y(t,\varepsilon)\leqq B$, $0\leqq t\leqq 1$, if $B\geqq A$; and if $Ae^{-1}\geqq B$,

$$Ae^{-t}-(Ae^{-1}-B)\exp[-Ae^{-1}(1-t)\varepsilon^{-1}]\leqq y(t,\varepsilon)\leqq Ae^{-t}+\varepsilon\gamma, \qquad 0\leqq t\leqq 1,$$

for a known positive constant γ. We next check for the occurrence of a Haber–Levinson crossing, i.e., we consider values of A, $B>0$ such that $\bar{u}_L(1)=Ae^{-1}<B$ and $u_R(0)=B<A$. Then it is easy to see that $\bar{u}_L(t_0)=u_R(t_0)=B$, for $t_0=\ln(AB^{-1})\in(0,1)$, with $\bar{u}'_L(t_0)=-B<u'_R(t_0)=0$. It only remains to observe that the crossing condition $f(t_0,\bar{u}_L(t_0),\omega)=-B\omega-\omega^2=-\omega(\omega+B)>0$ is satisfied for ω in $(-B,0)$. By the theorem of Haber and Levinson, for sufficiently small $\varepsilon>0$, (8.4), (8.5) has a unique solution $y=y(t,\varepsilon)$ satisfying

$$\lim_{\varepsilon\to0^+} y(t,\varepsilon)=\begin{cases}Ae^{-t}, & 0\leqq t\leqq \ln(AB^{-1}),\\ B, & \ln(AB^{-1})\leqq t\leqq 1.\end{cases}$$

Case 2. $A>0$, $B\leqq 0$. For these A and B, $\bar{u}_L(t)=Ae^{-t}$ and $\bar{u}_R(t)=Be^{1-t}$ are the stable roots of (8.6). Applying our boundary layer sign restrictions, i.e., $\bar{u}_L(1)=Ae^{-1}\geqq B$ and $\bar{u}_R(0)=Be\geqq A$, we find, since $B\leqq 0<A$, that the inequality $\bar{u}_L(1)\geqq B$ is always satisfied, while the inequality $\bar{u}_R(0)\geqq A$ never holds. Consequently, by Theorem 4.1, for each $\varepsilon>0$ sufficiently small, the problem (8.4), (8.5) has a solution $y=y(t,\varepsilon)$ such that

$$Ae^{-t}-(Ae^{-1}-B)\exp[-Ae^{-1}(1-t)\varepsilon^{-1}]\leqq y(t,\varepsilon)\leqq Ae^{-t}+\varepsilon\gamma, \qquad 0\leqq t\leqq 1,$$

that is,

$$\lim_{\varepsilon\to0^+} y(t,\varepsilon)=Ae^{-t}, \qquad 0\leqq t<1.$$

In this case there are no internal crossings.

Case 3. $A\leqq 0$, $B>0$. The stable roots of (8.6) are $u_L(t)=A$ and $u_R(t)=B$. Since $B>A$, Theorem 4.1 implies that for each sufficiently small $\varepsilon>0$, (8.4), (8.5) has a solution $y=y(t,\varepsilon)$ with

$$B-(B-A)\exp[-Bt\varepsilon^{-1}]\leqq y(t,\varepsilon)\leqq B, \qquad 0\leqq t\leqq 1,$$

that is,

$$\lim_{\varepsilon\to0^+} y(t,\varepsilon)=B, \qquad 0<t\leqq 1.$$

Again there are no internal crossings in this case.

Case 4. $A<0$, $B<0$. The stable roots of (8.6) are $u_L(t)=A$ and $\bar{u}_R(t)=Be^{1-t}$. In the case that $A\geqq B$, we apply Theorem 4.1 to see that a solution $y=y(t,\varepsilon)$ of (8.4), (8.5) exists and satisfies, for small $\varepsilon>0$,

$$A-(A-B)\exp[A(1-t)\varepsilon^{-1}]\leqq y(t,\varepsilon)\leqq A, \qquad 0\leqq t\leqq 1.$$

Similarly, if $\bar{u}_R(0)=Be\geqq A$, we have

$$Be^{1-t}-(Be-A)\exp[Bt\varepsilon^{-1}]-\varepsilon\gamma\leqq y(t,\varepsilon)\leqq Be^{1-t}, \qquad 0\leqq t\leqq 1.$$

Finally if $u_L(1)=A<B$ and $\bar{u}_R(0)=Be<A$, a Haber–Levinson crossing occurs at $t_0=\ln(BA^{-1}e)$.

It remains for us to determine the asymptotic behavior of solutions for the two cases: (i) $A<0, B=0$ and (ii) $A=0, B<0$. In the first case $u_L(t)=A$ is a stable root of (8.6); however, since $B=0$, $u_R=\bar{u}_R\equiv 0$ and $f_{y'}[0]=f_y[0]\equiv 0$. Since $u_L(1)=A<B$ and since u_L is never zero (i.e., does not intersect $u_R=\bar{u}_R$), this function does not determine the asymptotic behavior of solutions of (8.4), (8.5). We thus examine $u_R\equiv 0$ more closely. It is clear that $\beta\equiv 0$ is an upper solution of (8.4), (8.5), and a short computation shows that

$$\alpha(t)=\begin{cases}(A^{-1}-t\varepsilon^{-1/2})^{-1}-\gamma t\varepsilon^{1/2}, & 0\leqq t\leqq 1-\tau(\varepsilon),\\ (A^{-1}-(1-\tau(\varepsilon))\varepsilon^{-1/2})^{-1}-\gamma(1-\tau(\varepsilon))\varepsilon^{1/2}, & 1-\tau(\varepsilon)\leqq t\leqq 1,\end{cases}$$

is a lower solution, provided that $\gamma>0$ is appropriately chosen. Here $\tau(\varepsilon)>0$ is of order $O(\varepsilon^{1/2})$. Applying Theorem 2.2 we conclude that for $A<0, B=0$, (8.4), (8.5) has a solution $y=y(t,\varepsilon)$ for each $\varepsilon>0$ sufficiently small such that $(A^{-1}-t\varepsilon^{-1/2})^{-1}-\bar{\gamma}\varepsilon^{1/2}\leqq y(t,\varepsilon)\leqq 0, 0\leqq t\leqq 1$. In case (ii) one observes that the function $z(t,\varepsilon)=(B^{-1}-(2\varepsilon)^{-1}(1-t))^{-1}$ satisfies $z(0,\varepsilon)\leqq 0, z(1,\varepsilon)=B$, and $\varepsilon z''=-zz', 0<t<1$. Therefore, z is a lower solution, and $\beta\equiv 0$ is again an upper solution, and we deduce from Theorem 2.1 the existence of a solution $y=y(t,\varepsilon)$ satisfying

$$(B^{-1}-(2\varepsilon)^{-1}(1-t))^{-1}\leqq y(t,\varepsilon)\leqq 0, \qquad 0\leqq t\leqq 1.$$

These two cases illustrate the difficulty alluded to in § 4 regarding the determination of boundary layer behavior when appropriate partial derivatives are not strictly positive or strictly negative. One must treat such problems individually without recourse to a general theory.

Our results for Example 8.2 are summarized in the following boundary value portrait. (See Fig. 2.)

Example 8.3. The next example is chosen to illustrate how the existence of a singular solution of the reduced equation affects the asymptotic nature of solutions of the full problem for certain values of the boundary data. Consider the problem

$$\varepsilon y''=y-(y')^2, \qquad 0<t<1, \tag{8.7}$$

$$y(0,\varepsilon)=A, \qquad y(1,\varepsilon)=B. \tag{8.8}$$

The reduced equation is

$$0=u-(u')^2 \tag{8.9}$$

which has the regular solutions, defined for $A, B\geqq 0$,

$$u_L(t)=\tfrac{1}{4}(2A^{1/2}-t)^2, \qquad u_R(t)=\tfrac{1}{4}(t+2B^{1/2}-1)^2,$$

and

$$\bar{u}_L(t)=\tfrac{1}{2}(2A^{1/2}+t)^2, \qquad \bar{u}_R(t)=\tfrac{1}{4}(2B^{1/2}+1-t)^2,$$

as well as the singular solution $u_s\equiv 0$. To check the stability of these roots, set $f(t,y,y')=y-(y')^2$, then $f_{y'}=-2y'$ and $f_y\equiv 1$. It follows that $\bar{u}_L$ and $\bar{u}_R$ are

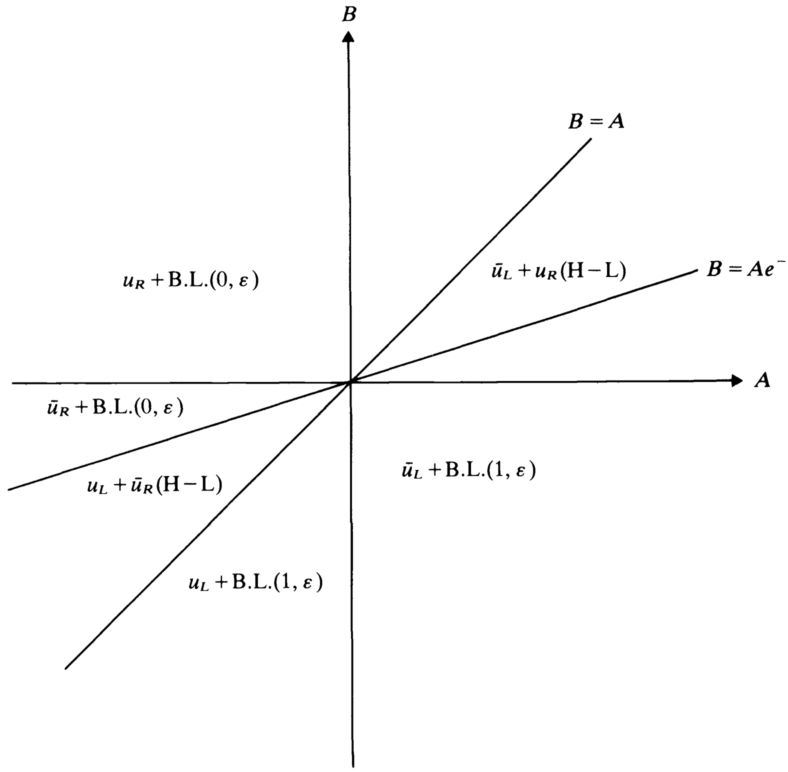

FIG. 2

unstable, while

$$f_{y'}[u_L(t)] = 2A^{1/2} - t\begin{cases} \geqq 0, & t \leqq 2A^{1/2}, \\ <0, & t > 2A^{1/2}, \end{cases}$$

$$f_{y'}[u_R(t)] = 1 - 2B^{1/2} - t\begin{cases} \leqq 0, & t \geqq 1 - 2B^{1/2}, \\ >0, & t < 1 - 2B^{1/2}. \end{cases}$$

Consequently, u_L and u_R are globally stable if $A^{1/2} \geqq \frac{1}{2}$ and $B^{1/2} \geqq \frac{1}{2}$, respectively. If $0 < A^{1/2}, B^{1/2} < \frac{1}{2}$, then these roots lose stability at $t_1 = 2A^{1/2}$, $t_2 = 1 - 2B^{1/2}$, respectively. Finally $u_s \equiv 0$ is globally stable since $f_y \equiv 1$.

Case 1. $A^{1/2}, B^{1/2} > \frac{1}{2}$. We check first for boundary layer behavior. Since $f_{y'y'} \equiv -2$, we require $u_L(1) = \frac{1}{4}(2A^{1/2} - 1)^2 \geqq B$ and $u_R(0) = \frac{1}{4}(2B^{1/2} - 1)^2 \geqq A$. For such A and B we apply Theorem 4.1 to conclude that for each $\varepsilon > 0$ sufficiently small, (8.7), (8.8) has a solution $y = y(t, \varepsilon)$ such that

$$\frac{1}{4}(2A^{1/2} - t)^2 - \left(\frac{1}{4}(2A^{1/2} - 1)^2 - B\right) \exp\left[-(2A^{1/2} - 1)(1 - t)\varepsilon^{-1}\right] \leqq y(t, \varepsilon)$$

$$\leqq \frac{1}{4}(2A^{1/2} - t)^2 + \frac{\varepsilon}{2}, \qquad 0 \leqq t \leqq 1,$$

if $\frac{1}{4}(2A^{1/2}-1)^2 \geqq B$; while if $\frac{1}{4}(2B^{1/2}-1)^2 \geqq A$,

$$\frac{1}{4}(t+2B^{1/2}-1)^2 - \left(\frac{1}{4}(2B^{1/2}-1)^2 - A\right) \exp\left[-(2B^{1/2}-1)t\varepsilon^{-1}\right] \leqq y(t,\varepsilon)$$
$$\leqq \frac{1}{4}(t+2B^{1/2}-1)^2 + \frac{\varepsilon}{2}, \qquad 0 \leqq t \leqq 1.$$

Next, if $u_L(1) = \frac{1}{4}(2A^{1/2}-1)^2 < B$ *and* $u_R(0) = \frac{1}{4}(2B^{1/2}-1)^2 < A$, it is easy to see that a Haber–Levinson crossing occurs at $t_0 = A^{1/2} - B^{1/2} + \frac{1}{2} \in (0, 1)$.

Case 2. $A, B \leqq 0$. For such A and B there are no regular solutions of (8.9) which satisfy either of the boundary conditions at $t = 0$ or $t = 1$. However, since $u_s \equiv 0$ is stable on $[0, 1]$ (with $f_y \equiv 1$) and $u_s(0) \geqq A$, $u_s(1) \geqq B$, we can apply Theorem 4.2 (with $q = 0$) to conclude that for all $\varepsilon > 0$, there exists a solution $y = y(t, \varepsilon)$ of (8.7), (8.8) satisfying

$$A \exp\left[-t\varepsilon^{-1/2}\right] + B \exp\left[-(1-t)\varepsilon^{-1/2}\right] \leqq y(t,\varepsilon) \leqq 0, \qquad 0 \leqq t \leqq 1.$$

Case 3. $B > 0$, $A \leqq 0$. In this case, $u_R(t) = \frac{1}{4}(t+2B^{1/2}-1)^2$ exists on $[0, 1]$; however, there is no regular solution of (8.9) satisfying the boundary condition at $t = 0$. As noted above, if $B^{1/2} > \frac{1}{2}$, then u_R is stable on $[0, 1]$, and since $u_R(0) \geqq A$, we can apply Theorem 4.1 to conclude that for all sufficiently small $\varepsilon > 0$, (8.7), (8.8) has a solution $y = y(t, \varepsilon)$ with

$$\frac{1}{4}(t+2B^{1/2}-1)^2 - \left(\frac{1}{4}(2B^{1/2}-1)^2 - A\right) \exp\left[-(2B^{1/2}-1)t\varepsilon^{-1}\right]$$
$$\leqq y(t,\varepsilon) \leqq \frac{1}{4}(t+2B^{1/2}-1)^2 + \frac{\varepsilon}{2}, \qquad 0 \leqq t \leqq 1.$$

If, on the other hand, $0 < B^{1/2} < \frac{1}{2}$, we know that u_R becomes unstable at $t_2 = 1 - 2B^{1/2} \in (0, 1)$. However, at this point, $u_R(t_2) = 0 = u_s(t_2)$, i.e., u_R crosses the singular solution $u_s \equiv 0$. Since $A \leqq 0$ and therefore $u_s(0) \geqq A$, we can apply Theorem 6.4 to conclude that for such A and B (and ε sufficiently small), (8.7), (8.8) has a solution $y = y(t, \varepsilon)$ satisfying

$$\lim_{\varepsilon \to 0^+} y(t,\varepsilon) = \begin{cases} 0, & 0 < t \leqq 1 - 2B^{1/2}, \\ \frac{1}{4}(t+2B^{1/2}-1)^2, & 1 - 2B^{1/2} \leqq t \leqq 1. \end{cases}$$

Finally if $B^{1/2} = \frac{1}{2}$, then $f_{y'}[u_R(t)] \leqq 0$, $0 \leqq t \leqq 1$; however, since $f_y \equiv 1$, we apply the appropriate modification of Theorem 4.2 to conclude that for sufficiently small $\varepsilon > 0$, (8.7), (8.8) has a solution satisfying

$$\frac{1}{4}t^2 + A \exp\left[-t\varepsilon^{-1/2}\right] \leqq y(t,\varepsilon) \leqq \frac{1}{4}t^2 + \frac{\varepsilon}{2}, \qquad 0 \leqq t \leqq 1.$$

Case 4. $A > 0$, $B \leqq 0$. This case is the reflection of Case 3; consequently, the statements made in Case 3 apply with B replaced by A, u_R replaced by u_L, and t replaced by $1 - t$. We omit the details, except to note that a regular-singular crossing occurs for $0 < A^{1/2} < \frac{1}{2}$ between u_L and $u_s \equiv 0$ at the point $t_1 = 2A^{1/2}$.

Case 5. $0 < A^{1/2}, B^{1/2} < \frac{1}{2}$. This is the final combination of A and B left to consider. As noted above, u_L becomes unstable at $t_1 = 2A^{1/2} \in (0, 1)$ and u_R

becomes unstable at $t_2=1-2B^{1/2}\in(0,1)$. In addition, $u_s(0)<A$ and $u_s(1)<B$; thus, there can be no solutions of (8.7), (8.8) for such A, B which possess boundary layers. We distinguish however two types of interior crossings:

(i) $A^{1/2}+B^{1/2}<\frac{1}{2}$. In this case $t_1<t_2$, but $u_L(t_1)=0$ and $u_R(t_2)=0$. By Theorem 6.1, for all sufficiently small $\varepsilon>0$, we conclude that (8.7), (8.8) has a solution $y=y(t,\varepsilon)$ satisfying

$$\lim_{\varepsilon\to 0^+} y(t,\varepsilon)=\begin{cases}\frac{1}{4}(2A^{1/2}-t)^2, & 0\leqq t\leqq 2A^{1/2},\\ 0, & 2A^{1/2}\leqq t\leqq 1-2B^{1/2},\\ \frac{1}{4}(t+2B^{1/2}-1)^2, & 1-2B^{1/2}\leqq t\leqq 1.\end{cases}$$

(ii) $A^{1/2}+B^{1/2}>\frac{1}{2}$. For these A and B, $t_1>t_2$ and it is easy to see that a Haber–Levinson crossing occurs at the point $t_0=A^{1/2}-B^{1/2}+\frac{1}{2}\in(t_2,t_1)$.

Finally if $A^{1/2}+B^{1/2}=\frac{1}{2}$, then $u_L\equiv u_R$. We summarize our discussion of this example in the form of a boundary value portrait. (See Fig. 3.)

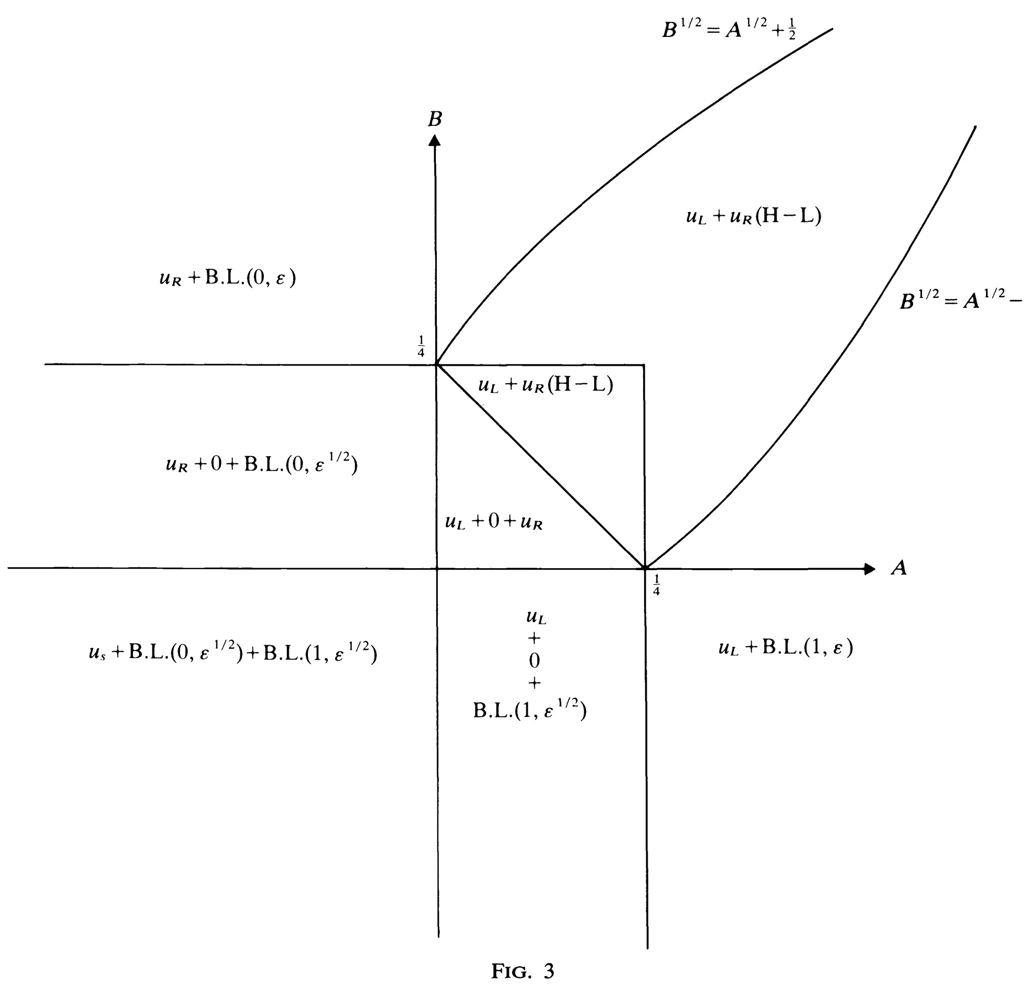

FIG. 3

Example 8.4. The final example of this section possesses solutions with algebraic boundary layer terms of the form discussed in Theorem 4.2. Also it illustrates why we restricted ourselves in Theorem 6.4 to functions $f=f(t, y, y')$ satisfying $f_y[u_s]>0$.

Consider the problem

$$\varepsilon y''=y^3-(y')^2, \qquad 0<t<1, \tag{8.10}$$

$$y(0, \varepsilon)=A, \qquad y(1, \varepsilon)=B. \tag{8.11}$$

The corresponding reduced equation

$$0=u^3-(u')^2 \tag{8.12}$$

has regular solutions of the form $u(t)=4(t+c)^{-2}$ and the singular solution $u_s\equiv 0$. Clearly the functions $u_L(t)=4(t+2A^{-1/2})^{-2}$ and $u_R(t)=4(2B^{-1/2}+1-t)^{-2}$ are stable roots of (8.12) for all positive A and B.

Checking first for boundary layer behavior, we require, since $f_{y'y'}\equiv -2(f(t, y, y')=y^3-(y')^2)$, $u_L(1)\geqq B$ and $u_R(0)\geqq A$, i.e., $\frac{1}{2}+A^{-1/2}\leqq B^{-1/2}$ and $\frac{1}{2}+B^{-1/2}\leqq A^{-1/2}$, respectively. For such $A, B>0$, we apply Theorem 4.1 to conclude that for all sufficiently small $\varepsilon>0$, (8.10), (8.11) has a solution $y=y(t, \varepsilon)$ such that

$$4(t+2A^{-1/2})^{-2}-(4(1+2A^{-1/2})^{-2}-B)\exp[-k_1(1-t)\varepsilon^{-1}]\leqq y(t, \varepsilon)$$
$$\leqq 4(t+2A^{-1/2})^{-2}+\varepsilon\gamma, \qquad 0\leqq t\leqq 1,$$

if $\frac{1}{2}+A^{-1/2}\leqq B^{-1/2}$; while if $\frac{1}{2}+B^{-1/2}\leqq A^{-1/2}$,

$$4(2B^{-1/2}+1-t)^{-2}-(4(2B^{-1/2}+1)^{-2}-A)\exp[-k_2 t\varepsilon^{-1}]$$
$$\leqq y(t, \varepsilon)\leqq 4(2B^{-1/2}+1-t)^{-2}+\varepsilon\gamma, \qquad 0\leqq t\leqq 1.$$

Here $k_1=16(1+2A^{-1/2})^{-3}$, $k_2=16(1+2B^{-1/2})^{-3}$ and γ is a positive constant.

Next, if $u_L(1)<B$ and $u_R(0)<A\,(A, B>0)$, it is easy to see that a Haber–Levinson crossing occurs at $t_0=\frac{1}{2}+B^{-1/2}-A^{-1/2}\in(0, 1)$.

Suppose now that $B>0$ and $A\leqq 0$; then trivially $u_R(0)=4(1+2B^{-1/2})^{-2}\geqq A$. Again it follows from Theorem 4.1 that a solution $y=y(t, \varepsilon)$ of (8.10), (8.11) exists and satisfies $\lim_{\varepsilon\to 0^+} y(t, \varepsilon)=4(2B^{-1/2}+1-t)^{-2}$, $0<t\leqq 1$. Similarly, if $A>0$ and $B\leqq 0$, $u_L(1)=4(1+2A^{-1/2})\geqq B$, and so there exists a solution $y=y(t, \varepsilon)$ such that $\lim_{\varepsilon\to 0^+} y(t, \varepsilon)=4(t+2A^{-1/2})^{-2}$, $0\leqq t<1$.

It only remains to determine the asymptotic behavior when $A\leqq 0$ and $B\leqq 0$. The singular solution $u_s\equiv 0$ evidently satisfies $u_s(0)\geqq A$ and $u_s(1)\geqq B$. Thus we can apply Theorem 4.2 (with $q=1$) to deduce that for all $\varepsilon>0$, (8.10), (8.11) has a solution $y=y(t, \varepsilon)$ such that

$$A(1-A(2\varepsilon)^{-1/2}t)^{-1}+B(1-B(2\varepsilon)^{-1/2}(1-t))^{-1}\leqq y(t, \varepsilon)\leqq 0, \qquad 0\leqq t\leqq 1.$$

We note that for this example there are no crossings between u_L or u_R and $u_s\equiv 0$, i.e., u_L and u_R have no real zeros. Our results are summarized in the following boundary value portrait. (See Fig. 4.)

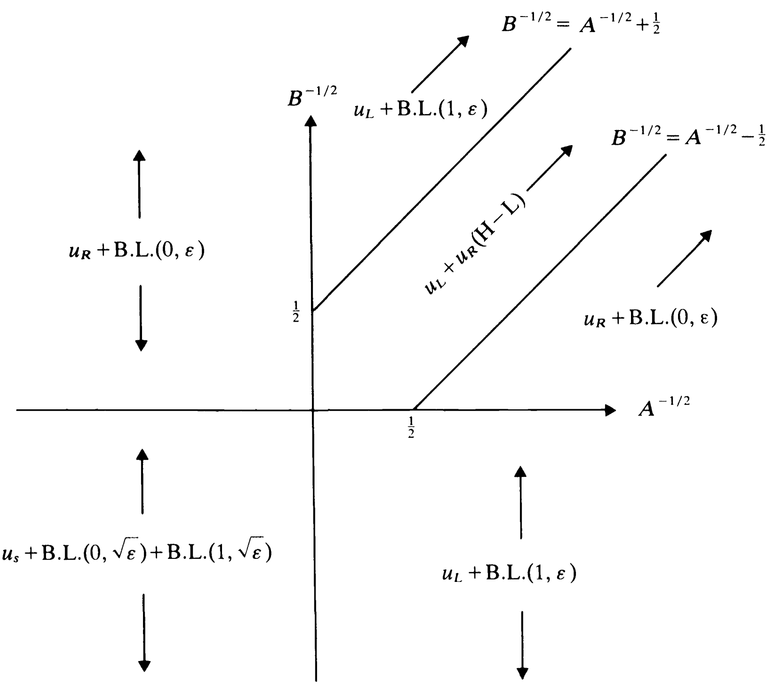

FIG. 4

9. Properties of solutions. We briefly discuss here the behavior of solutions of problems to which none of our hypotheses of the previous sections applies. We first indicate what types of behavior are excluded by these hypotheses, in particular, by our restrictions on $f_{y'y'}$.

As noted in § 4, the assumption that $f_{y'y'}$ never vanishes in the domain of interest places a restriction on the convexity properties which solutions possess inside boundary layer regions. Essentially if $f_{y'y'} < 0$, then inside such a region, any solution y is concave since y'' is negative there; while if $f_{y'y'} > 0$, solutions are convex. Equally important, such sign restrictions prevent the occurrence of interior transition layers (i.e., shock layers) in the solution y itself. The reason for this is simply that across a shock layer (see Fig. 5, below) the sign of y'' changes rapidly for small $\varepsilon > 0$, but such sign changes are impossible if $f_{y'y'} \neq 0$. A similar line of reasoning shows that solutions y cannot exhibit densely oscillatory behavior if $f_{y'y'} \neq 0$ since such oscillations are possible only if y'' changes sign infinitely often.

The convexity properties of solutions inside of boundary layers also determine the existence or nonexistence of solutions for certain choices of the boundary conditions, as the following example shows.

Example 9.1. Consider the problem

$$\varepsilon y'' = y^2 + (y')^2, \qquad 0 < t < 1, \tag{9.1}$$

$$y(0, \varepsilon) = A, \qquad y(1, \varepsilon) = B. \tag{9.2}$$

The reduced equation $u^2 + (u')^2 = 0$ clearly has $u \equiv 0$ as its only real solution.

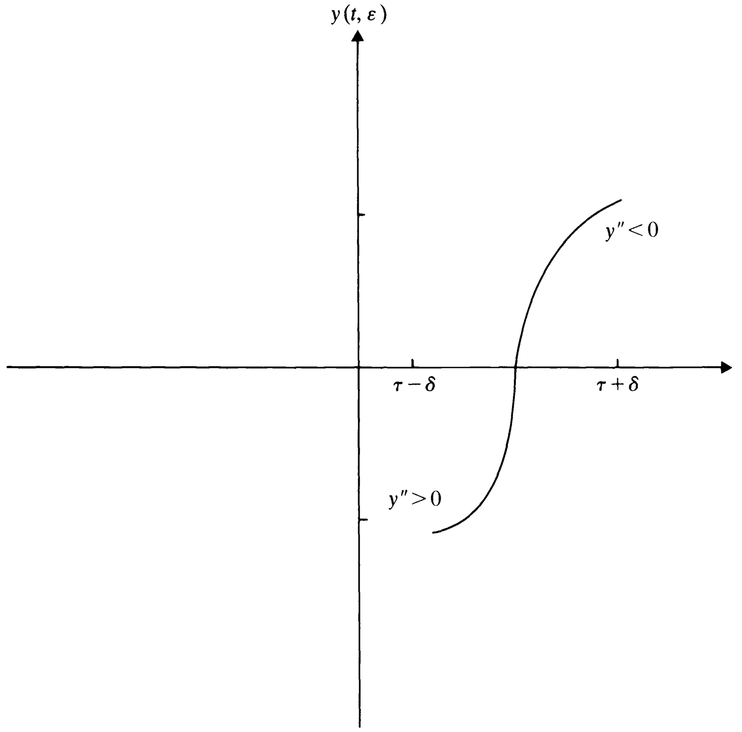

FIG. 5

Consequently, if A and B are nonnegative we can apply Theorem 4.3 (with $l=1$) to conclude that for all $\varepsilon>0$, (9.1), (9.2) has a solution $y=y(t,\varepsilon)$ with

$$0 \leqq y(t,\varepsilon) \leqq A(1+(A(6\varepsilon)^{-1})^{1/2}t)^{-2}+B(1+(B(6\varepsilon)^{-1})^{1/2}(1-t))^{-2}, \qquad 0 \leqq t \leqq 1.$$

It turns out that for all other values of A and B none of the previous theory applies. However, for these A and B we argue as follows. Suppose, for example, that A is negative. Then a solution y of (9.1), (9.2) for sufficiently small $\varepsilon>0$ must pass through the point $(0, A)$ and must reach an order $O(\varepsilon)$-neighborhood of the reduced solution $u\equiv 0$ in a time interval, $0\leqq t\leqq \delta$, of length $O(\varepsilon)$. But since $A<0$, either $y''(t)<0$, $0\leqq t\leqq\delta$, or $y''(t)>0$, $0\leqq t\leqq \delta_1$, and $y''(t)<0$, $\delta_2\leqq t\leqq\delta$. (See Figs. 6a, 6b.) However, in $[0,\delta]$, $y''>0$, and so neither of these situations can occur. We conclude that (9.1), (9.2) has no solution if $A<0$ and ε is sufficiently small. A similar argument at $t=1$ shows the nonexistence of solutions if $B<0$.

We consider now a subclass of the problems we have been considering to which some of our assumptions do not apply, and yet, the solution exists and behaves in a regular manner for small $\varepsilon>0$. These problems are distinguished by the property that certain regular solutions u of the reduced equation $f(t,u,u')=0$ i.e., those for which $f_{y'}[u]\not\equiv 0$, behave like singular solutions in that one cannot impose *any* boundary condition on them. For such solutions our previous stability restrictions must be modified.

Recall that a regular reduced root which satisfied the right-hand boundary condition, i.e., u_R, was termed stable if $f_{y'}[u_R]$ was nonpositive on $[t_R, 1)$, while a

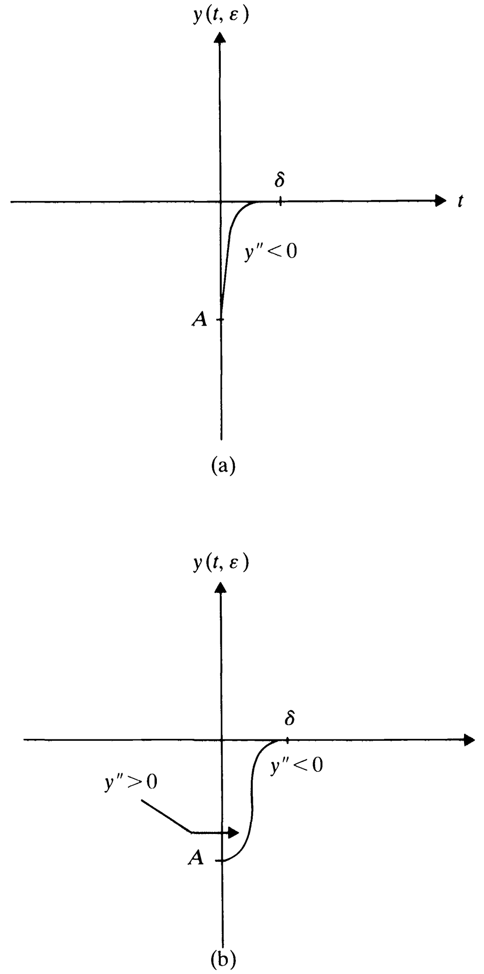

FIG. 6

root u_L (for which $u_L(0)=y(0, \varepsilon)$) was termed stable if $f_{y'}[u_L(t)]$ was nonnegative on $[0, t_L]$. However, if a regular root u cannot be made to satisfy either a condition at $t=0$ or $t=1$, then we define stability as follows: if $u_L(0)\neq y(0, \varepsilon)$, then $f_{y'}[u(t)]<0$, $0\leqq t\leqq \delta$, for a positive constant δ; while if $u(1)\neq y(1, \varepsilon)$, then $f_{y'}[u(t)]>0$, $1-\delta\leqq t\leqq 1$. In the rest of the interval $(\delta, 1-\delta)$, the signs of these functions are immaterial provided that u has some form of y-stability, e.g., $f_y[u(t)]>0$, $\delta\leqq t\leqq 1-\delta$. As regards boundary layer behavior, these restrictions on $f_{y'}[u]$ are very natural, since in all of our previous work, the crucial assumptions were that, near $t=0$, $f_{y'}[u(t)]\leqq -k<0$ (if $u(0)\neq y(0, \varepsilon)$), and $f_{y'}[u(t)]\geqq k>0$, near $t=1$ (if $u(1)\neq y(1, \varepsilon)$).

In the solution of an actual problem, one must then look for these "regular-singular" reduced roots and check their stability according to the above criteria. These stable roots can thus be included in the algorithm given above to the extent that they may have Haber–Levinson crossings with other regular roots or with each other, in addition to boundary layer behavior at one or both end-points.

As an illustration of the type of result which is possible for this class of problems, we give the following theorem. A proof can be found in [10]; however, the reader can easily provide one based on the calculations given above.

THEOREM 9.1. *Assume*

1) *the reduced equation* $f(t, u, u')=0$ *has a solution* $u=u(t)$ *of class* $C^{(2)}[a, b]$;

2) *the function* f *is continuous in* (t, y, y') *and of class* $C^{(2)}$ *with respect to* y, y' *in* E: $a \leqq t \leqq b$, $|y-u(t)| \leqq d(t, \varepsilon)$, $|y'|<\infty$, *where* $d(t, \varepsilon) \geqq |A-u(a)|$ (*if* $u(a) \neq A$), $a \leqq t \leqq a+\delta$, $d(t, \varepsilon) \geqq |B-u(b)|$(*if* $u(b) \neq B$), $b-\delta \leqq t \leqq b$, *and* $d(t, \varepsilon)=O(\varepsilon)$, $a+\delta<t<b-\delta$; *here* $\delta>0$ *is a small constant*; *also*, $f_{y'y'}=O(1)$ *in* E;

3) *if* $u(a) \neq A$, *then there exists a positive constant* k_1 *such that* $f_{y'}(t, u(t), u'(t)) \leqq -k_1<0$, *in* $E \cap [a, a+\delta]$; *if* $u(b) \neq B$, *then there exists a positive constant* k_2 *such that* $f_{y'}(t, u(t), u'(t)) \geqq k_2>0$, *in* $E \cap [b-\delta, b]$; *moreover, there exists a positive constant* p *such that*

$$(A-u(a))f_{y'y'} \geqq p>0, \quad in\, E \cap [a, a+\delta],$$

if $u(a) \neq A$, *and*

$$(B-u(b))f_{y'y'} \geqq p>0, \quad in\, E \cap [b-\delta, b],$$

if $u(b) \neq B$;

4) *there exists a positive constant* l *such that*

$$f_y(t, u(t), u'(t)) \geqq l>0, |y-u(t)| \leqq d(t, \varepsilon), a+\delta \leqq t \leqq b-\delta.$$

Then for each $\varepsilon>0$, ε *sufficiently small, the problem*

$$\varepsilon y''=f(t, y, y'), \qquad a<t<b,$$
$$y(a, \varepsilon)=A, \qquad y(b, \varepsilon)=B,$$

has a solution $y=y(t, \varepsilon)$. *In addition, for* t *in* $[a, b]$,

$$y(t, \varepsilon)=u(t)+O(|A-u(a)| \exp[-k_1(t-a)\varepsilon^{-1}]) \\ +O(|B-u(b)| \exp[-k_2(b-t)\varepsilon^{-1}])+O(\varepsilon).$$

We remark that this result can be viewed as a generalization of Theorem 4 in [2]. A simple illustration is contained in the next example.

Example 9.2. Consider the problem

$$\varepsilon y''=y-2ty'+(y')^2, \qquad -1<t<1, \tag{9.3}$$

$$y(-1, \varepsilon)=A, \qquad y(1, \varepsilon)=B. \tag{9.4}$$

The corresponding reduced equation

$$u=2tu'-(u')^2, \qquad -1<t<1,$$

has the regular-singular roots $u_1(t)=\frac{3}{4}t^2$ and $u_2(t)\equiv 0$. Of these, u_1 is stable in the sense outlined above, since

$$f_{y'}[u_1(t)] = t\begin{cases} <0, & -1 \leqq t \leqq -1+\delta, \\ >0, & 1-\delta \leqq t \leqq 1, \end{cases} \quad \text{for } 0<\delta<1.$$

Since $f_{y'y'} \equiv 2$, we can apply Theorem 9.1 if A, B satisfy $A, B \geqq \frac{3}{4}$ to deduce that (9.3), (9.4) has a (unique) solution $y = y(t, \varepsilon)$ for which

$$\tfrac{3}{4}t^2 \leqq y(t, \varepsilon) \leqq \tfrac{3}{4}t^2 + (A - \tfrac{3}{4}) \exp[-k(1+t)\varepsilon^{-1}] + (B - \tfrac{3}{4}) \exp[-k(1-t)\varepsilon^{-1}] + 2\varepsilon, \qquad -1 \leqq t \leqq 1, \quad \text{where } 0<k<1.$$

The existence of such a root u_1 is intimately connected with the fact that the reduced equation has a singular point at $t=0$ across which $f_{y'}[u_1(t)]$ changes from negative to positive.

Finally we note that for other choices of the boundary conditions, the stable root u_1 is involved in Haber–Levinson crossings with other regular roots and also in a smooth crossing with $u_2 \equiv 0$ at $t=0$. For a description of some of these phenomena, see [11] and [12].

10. Concluding remarks. The present study of the boundary value problem (7.1), (7.2) has been based upon the principal assumptions that the partial derivative $f_{y'y'}$ is never zero and that the corresponding reduced equation has smooth solutions with certain stability properties. It is a natural question to ask how the existence and asymptotic behavior of solutions are affected if either (or both) of these assumptions is altered. In particular, if $f_{y'y'}$ is allowed to vanish at certain values of t or along certain roots of the reduced equation, then the nature of the sign change is critical in determining asymptotic behavior. Among other things solutions can possess shock layers and there can exist "angular" regular–singular crossings. On the other hand, if the smoothness requirement on the reduced solutions is relaxed, solutions of the full problem can possess singular layers (cf. [14, § 21]) associated with discontinuities in the derivatives of reduced roots. In addition, the basic result of Haber and Levinson given in § 5 must also be modified to handle such cases. These and other related problems are studied in several papers of the author [10], [11] and [12].

Acknowledgment. The author wishes to thank Professor J. B. Keller for inviting him to spend a year at the Courant Institute, and Professor W. R. Wasow for inviting him to spend an equally enjoyable year at Wisconsin. Thanks are also due to the National Science Foundation for their generous support of this and related efforts.

REFERENCES

[1] YU. P. BOGLAEV, *The two-point problem for a class of ordinary differential equations with a small parameter coefficient of the derivative*, USSR Computational Math. and Math. Phys., 10 (1970), no. 4, pp. 191–204.

[2] N. I. BRIŠ, *On boundary value problems for the equation $\varepsilon y'' = f(x, y, y')$ for small ε*, Dokl. Akad. Nauk SSSR, 95 (1954), pp. 429–432.

[3] D. COHEN, *Perturbation Theory*, Lecture Notes, AMS–SIAM Summer Institute on Continuum Modelling, July 1975.

[4] F. W. DORR, S. V. PARTER AND L. F. SHAMPINE, *Applications of the maximum principle to singular perturbation problems*, SIAM Rev., 15 (1973), pp. 43–88.

[5] J. E. FLAHERTY AND R. E. O'MALLEY, JR., *The numerical solution of boundary value problems for stiff differential equations*, Math. Comput., 31 (1977), pp. 66–93.

[6] S. HABER AND N. LEVINSON, *A boundary value problem for a singularly perturbed differential equation*, Proc. Amer. Math. Soc., 6 (1955), pp. 866–872.

[7] P. HABETS AND M. LALOY, *Etude de problemes aux limites par la methode des sur-et-sous solutions*, Lecture Notes, Catholic University of Louvain, Louvain, Belgium, 1974.

[8] F. A. HOWES, *Singular perturbations and differential inequalities*, Mem. Amer. Math. Soc., 168 (1976).

[9] ———, *A class of boundary value problems whose solutions possess angular limiting behavior*, Rocky Mountain J. Math., 6 (1976), pp. 591–607.

[10] ———, *Singularly perturbed nonlinear boundary value problems with turning points. II*, this Journal, 9 (1978), pp. 250–270.

[11] ———, *Modified Haber–Levinson crossings*, Trans. Amer. Math. Soc., to appear.

[12] ———, *A boundary layer theory for some linear and nonlinear boundary value problems*, Rocky Mountain J. Math., to appear.

[13] L. K. JACKSON, *Subfunctions and second-order ordinary differential inequalities*, Advances in Math., 2 (1968), pp. 307–363.

[14] R. E. MEYER, *Introduction to Mathematical Fluid Dynamics*, Wiley-Interscience, New York, 1971.

[15] M. NAGUMO, *Über die Differentialgleichung* $y''=f(x, y, y')$, Proc. Phys. Math. Soc. Japan, 19 (1937), pp. 861–866.

[16] R. E. O'MALLEY, JR., *On singular perturbation problems with interior nonuniformities*, J. Math. Mech., 19 (1970), pp. 1103–1112.

[17] ———, *Introduction to Singular Perturbations*, Academic Press, New York, 1974.

[18] ———, *Boundary layer methods for ordinary differential equations with small coefficients multiplying the highest derivatives*, Proc. Symp. on Constructive and Computational Methods for Differential and Integral Equations, Springer-Verlag Lecture Notes in Math., 430, Springer-Verlag, New York, 1974, pp. 363–389.

[19] A. B. VASIL'EVA, *Asymptotic behavior of solutions of certain problems involving nonlinear differential equations containing a small parameter multiplying the highest derivatives*, Russian Math. Surveys, 18 (1963), pp. 13–84.

[20] W. WASOW, *Asymptotic Expansions for Ordinary Differential Equations*, Wiley-Interscience, New York, 1965.

[21] ———, *The capriciousness of singular perturbations*, Nieuw Arch. Wisk., 18 (1970), pp. 190–210.

SIAM J. MATH. ANAL.
Vol. 9, No. 2, April 1978
Copyright © 1978 Society for Industrial and Applied Mathematics

SINGULARLY PERTURBED NONLINEAR BOUNDARY VALUE PROBLEMS WITH TURNING POINTS. II*

F. A. HOWES†

Abstract. The existence and the asymptotic behavior as $\varepsilon \to 0^+$ of solutions of the boundary value problem $\varepsilon y'' = f(t, y, y', \varepsilon)$, $-1 < t < 1$, $y(\pm 1, \varepsilon)$ prescribed, are studied in the case that $f_{y'y'} = O(1)$, as $|y'| \to \infty$ and $f_{y'y'}$ vanishes at $t = 0$. For small values of the perturbation parameter $\varepsilon > 0$, solutions are closely approximated by certain roots of the reduced equation $f(t, u, u', 0) = 0$ throughout most of the t-interval with the possible exception of a shock layer region centered at $t = 0$ and/or a boundary layer region at $t = -1$ or $t = 1$ (or both endpoints). Inside such regions a solution changes rapidly either to transfer from one reduced root to another or to satisfy the given boundary data. These and other related phenomena are illustrated by many examples.

1. Introduction. In this paper we continue our study of the existence and the asymptotic behavior as $\varepsilon \to 0^+$ of solutions of the singularly perturbed boundary value problem

$$\varepsilon y'' = f(t, y, y', \varepsilon), \qquad -1 < t < 1, \tag{1.1}$$

$$y(-1, \varepsilon) = A, \qquad y(1, \varepsilon) = B, \tag{1.2}$$

in the case when f possesses a turning point at $t = 0$. For linear functions f many results are known and the outline of a general theory can be found, for example, in [18, Chap. 8], [13] and [15, Chap. 8]. It is also possible to treat quasilinear functions f (i.e., affine functions of the derivative y'); some results are given in [3], and a more general approach is presented in [7]. However there does not seem to have been any treatment whatsoever of functions f which are quadratically nonlinear in y' and which, in a sense to be made precise below, possess a turning point at an interior point of $[-1, 1]$.

Consider then the boundary value problem (1.1), (1.2) in which f is a continuous function of its arguments and has continuous partial derivatives of the second order with respect to y and y'. Assume, in addition, that $\partial^2 f/\partial y'\partial y' = f_{y'y'}$ is of order $O(1)$, as $|y'| \to \infty$, and that $f_{y'y'} \not\equiv 0$. Then if $f_{y'y'}$ vanishes at $t = 0$, we say that f possesses a generalized turning point at $t = 0$. A word of explanation is perhaps appropriate. In the case of a linear function f, a turning point is defined (see, e.g., [18, Chap. 8] or [13]) as a point at which the coefficient of y' vanishes. (Of course, this coefficient is assumed not to vanish identically.) In [7] we extended this definition to nonlinear functions f by requiring that the partial derivative $\partial f/\partial y' = f_{y'}$ vanishes at $t = 0$, in a rather large domain of variation of y and y'. However in dealing with functions f which are quadratically nonlinear in y', it is necessary to restrict ourselves essentially to evaluating $f_{y'}$ along a single path $u = u(t)$ which turns out to be a solution of the corresponding reduced equation $f(t, u, u', 0) = 0$. Consequently, the results in [7] are, for the most part, inapplicable to the more nonlinear functions considered here, although the theorems

* Received by the editors March 11, 1976, and in revised form August 6, 1976.

† Mathematics Department, University of Wisconsin—Madison, Madison, Wisconsin. Now at School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455. This work was supported in part by the Mathematics Department University of Wisconsin—Madison.

presented below share an affinity with those of [7] both in regard to formulation and method of proof.

We restrict ourselves to functions f which are quadratically nonlinear in y' in order to study solutions of (1.1), (1.2) which possess boundary layer and interior transition layer behavior. The classic result of Vishik and Liusternik [17] (see also the discussion in [16, Chap. 2]) naturally limits such behavior to solutions of problems whose righthand sides grow no faster thay $(y')^2$, as $|y'|\to\infty$. In addition, the method of proof requires such a growth restriction, since as in [7], we employ the theorem of Nagumo–Jackson [11], [12] as well as an extended version of Habets and Laloy [6].

Before commencing our study of the problem (1.1), (1.2) in which f has a generalized turning point, we make a few remarks about the previous work on problems whose righthand sides depend on $(y')^2$. The first result seems to be the classic theorem of Haber and Levinson [5] (see also [16, Chap. 2] and [14] for later extensions) which of course treats functions f without restriction to quadratic nonlinearities in y'. It describes the asymptotic behavior which can occur when two stable solutions of the reduced equation (each satisfying the boundary condition at the appropriate endpoint) intersect with unequal slopes at an interior point of the t-interval. The second treatment of such problems is contained in the paper of Dorr, Parter and Shampine [4, § 5]. These authors considered the possibility that solutions of the boundary value problem

$$\varepsilon y''+p(t,y)(y')^2+g(t,y)y'-b(t,y)y=F(t,y),\qquad 0<t<1,$$
$$y(0,\varepsilon)=A,\quad y(1,\varepsilon)=B,\quad A<B,\quad B>0,$$

where $p(t,y)<0$ for all t and y, may exhibit boundary layer behavior as well as the "Haber–Levinson" behavior just described. Finally in [9] we have considered the general problem (1.1), (1.2) under the principal assumption that $f_{y'y'}\neq 0$. Our results amplify and extend those of [4, § 5]; in addition, some new phenomena involving interior crossings of reduced solutions are treated for the first time. The theorems presented below complement several of those in [9]; however, the types of behavior which can occur when $f_{y'y'}$ has a zero in $(-1,1)$ are sufficiently distinctive and varied to warrant the special treatment given here.

We present briefly in the next section the extension of the classical Nagumo theory which is required in studying the problem (1.1), (1.2). In the following sections we isolate the various types of asymptotic behavior which solutions can exhibit and we motivate the hypotheses of the theorems with some illustrative examples.

2. A differential inequality theorem. In the context of the problem (1.1), (1.2), the classic theorem of Nagumo [12] (see Jackson's paper [11] for a more modern version) states that if there exist functions α,β of class $C^{(2)}[-1,1]$ with $\alpha\leqq\beta$, $\alpha(-1,\varepsilon)\leqq A\leqq\beta(-1,\varepsilon)$, $\alpha(1,\varepsilon)\leqq B\leqq\beta(1,\varepsilon)$, and for t in $(-1,1)$, $\varepsilon\alpha''\geqq f(t,\alpha,\alpha',\varepsilon)$, $\varepsilon\beta''\leqq f(t,\beta,\beta',\varepsilon)$, then the problem (1.1), (1.2) has a solution $y=y(t,\varepsilon)$ for such $\varepsilon>0$ with $\alpha(t,\varepsilon)\leqq y(t,\varepsilon)\leqq\beta(t,\varepsilon)$, $-1\leqq t\leqq 1$, provided that $f=O(|y'|^2)$, as $|y'|\to\infty$. For our study of this problem we require the following extension of this result due to Habets and Laloy [6].

THEOREM 2.1. *Suppose the function* $f=f(t, y, y', \varepsilon)$ *is continuous on* $[-1, 1]\times \mathbb{R}^2\times[0, \varepsilon_0]$ $(\varepsilon_0>0)$ *and satisfies the growth estimate* $f=O(|y'|^2)$, *as* $|y'|\to\infty$. *Suppose also that there exist continuous functions* α, β $(\alpha\leqq\beta)$ *which are piecewise-*$C^{(2)}$ *bounding solutions on* $[-1, 1]$, *i.e., there is a finite partition* $\{t_i\}$, $1\leqq i\leqq n$, *of* $[-1, 1]$ *such that in each subinterval* (t_i, t_{i+1}), α *and* β *are of class* $C^{(2)}$, *and for* $0<\varepsilon\leqq\varepsilon_1\leqq\varepsilon_0$, $\alpha(-1, \varepsilon)\leqq A\leqq\beta(-1, \varepsilon)$, $\alpha(1, \varepsilon)\leqq B\leqq\beta(1, \varepsilon)$, *and for* t *in* (t_i, t_{i+1}), $\varepsilon\alpha''\geqq f(t, \alpha, \alpha', \varepsilon)$, $\varepsilon\beta''\leqq f(t, \beta, \beta', \varepsilon)$, *with* $D_l\alpha(t_i, \varepsilon)\leqq D_r\alpha(t_i, \varepsilon)$ *and* $D_l\beta(t_i, \varepsilon)\geqq D_r\beta(t_i, \varepsilon)$, $1<i<n$. *(Here* D_l, D_r *denotes lefthand, respectively righthand, differentiation.)*

Then the boundary value problem $\varepsilon y''=f(t, y, y', \varepsilon)$, $y(-1, \varepsilon)=A$, $y(1, \varepsilon)=B$, *has for each* ε, $0<\varepsilon\leqq\varepsilon_1$, *a solution* $y=y(t, \varepsilon)$ *satisfying* $\alpha(t, \varepsilon)\leqq y(t, \varepsilon)\leqq \beta(t, \varepsilon)$, $-1\leqq t\leqq 1$.

Theorem 2.1 extends the theorem of Nagumo–Jackson described above by allowing the bounding functions α and β to have finitely many "corners", provided that the correct inequalities are satisfied at the corner points t_i. Its proof is an easy modification of the one given in [11] if one notes that the maximum of a finite number of lower solutions (i.e., functions satisfying the α-differential inequalities) is also a lower solution, while the minimum of a finite number of upper solutions (i.e., functions satisfying the β-differential inequalities) is an upper solution. We remark finally that the solution of (1.1), (1.2) is unique if $\partial f/\partial y=f_y$ is strictly positive for all values of its arguments, as follows from a direct application of the maximum principle.

3. Interior transition (shock) layers. We begin our study of the problem

$$\varepsilon y''=f(t, y, y', \varepsilon), \qquad -1<t<1, \tag{3.1}$$

$$y(-1, \varepsilon)=A, \qquad y(1, \varepsilon)=B, \tag{3.2}$$

by considering conditions on the function f which guarantee the existence of solutions possessing transition layer (or shock layer) behavior at $t=0$. Our first requirement is that the corresponding reduced equation

$$f(t, u, u', 0)=0 \tag{3.3}$$

has two solutions (roots) $u=u_L(t)$ and $u=u_R(t)$, defined and sufficiently smooth on $[-1, 0]$ and $[0, 1]$, respectively, which satisfy $u_L(-1)=A$, $u_R(1)=B$ and $u_L(0)\neq u_R(0)$. Then if these reduced roots are stable in the sense of assumption 5) of Theorem 3.1 below and if $f_{y'y'}$ satisfies the inequalities: $(u_L(0)-u_R(0))f_{y'y'}\leqq 0$, $-\delta\leqq t\leqq 0$, and $(u_L(0)-u_R(0))f_{y'y'}\geqq 0$, $0\leqq t\leqq\delta$, for some constant $\delta>0$, we anticipate that a solution $y=y(t, \varepsilon)$ of (3.1), (3.2) exists for small $\varepsilon>0$ and satisfies $y(t, \varepsilon)\to u_L(t)$, $-1\leqq t<0$, and $y(t, \varepsilon)\to u_R(t)$, $0<t\leqq 1$, as $\varepsilon\to 0^+$. The requirement that the reduced roots u_L and u_R should be stable is a natural and standard assumption (see, e.g., [16, Chap. 1] or [9]). However, the sign restrictions on $f_{y'y'}$ near $t=0$ deserve a brief comment. Since across a shock layer a solution $y(t, \varepsilon)$ of (3.1), (3.2) changes rapidly from being convex to concave (or vice versa), the reduced roots and the function $f_{y'y'}$ must reflect this change in convexity in a neighborhood of $t=0$. For instance, suppose $u_L(0)>u_R(0)$, then if $y(t, \varepsilon)$ passes through a shock layer at $t=0$, y changes from being concave to being convex in an order $o(1)$-neighborhood of $t=0$. In such a neighborhood, $y'(t, \varepsilon)$ is unbounded as a function of ε, and since $\varepsilon y'' \cong \frac{1}{2} f_{y'y'}(y')^2$ there, the sign of $f_{y'y'}$ must be compatible with this change in convexity, i.e., $tf_{y'y'} \geqq 0$. These heuristics are made precise in the following theorem.

THEOREM 3.1. *Assume*

1) *the reduced equation* (3.3) *has two solutions* $u = u_L(t)$, $u = u_R(t)$, *defined and of class* $C^{(2)}$ *on the respective intervals* $[-1, 0]$, $[0, 1]$, *with* $u_L(0) \neq u_R(0)$;

2) *the function f is continuous in* (t, y, y', ε) *and of class* $C^{(2)}$ *with respect to* y, y' *in* $\mathcal{R}(-1,1)$: $-1 \leqq t \leqq 1$, $|y - u_L(t)| \leqq d(t)$, $-1 \leqq t \leqq 0$, $|y - u_R(t)| \leqq d(t)$, $0 \leqq t \leqq 1$, *with* $d(t) = O(\varepsilon^{1/2})$, *for t in* $[-1, 1] \backslash (-\delta, \delta)$ *and* $d(t) \geqq |u_L(0) - u_R(0)|$, $-\delta < t < \delta$, *for* $\delta > 0$ *a small constant independent of* ε, $|y'| < \infty$, $0 \leqq \varepsilon \leqq \varepsilon_0$, $\varepsilon_0 > 0$; *and in* $\mathcal{R}(-1, 1)$, $f = O(|y'|^2)$, *as* $|y'| \to \infty$;

3) *for the same constant* $\delta > 0$ *in* 2), $(u_L(0) - u_R(0)) f_{y'y'} \leqq 0$ *in* $\mathcal{R}(-\delta, 0)$[1] *and* $(u_L(0) - u_R(0)) f_{y'y'} \geqq 0$ *in* $\mathcal{R}(0, \delta)$;

4) *there is a positive constant l such that*

$$f_y(t, y, u_L'(t), \varepsilon) \geqq l > 0, \qquad |y - u_L(t)| \leqq d(t), \qquad -1 \leqq t \leqq 0;$$
$$f_y(t, y, u_R'(t), \varepsilon) \geqq l > 0, \qquad |y - u_R(t)| \leqq d(t), \qquad 0 \leqq t \leqq 1;$$

5) $f_{y'}(t, y, u_L'(t), \varepsilon) \geqq h_1(t)$, $|y - u_L(t)| \leqq d(t)$, $-1 \leqq t \leqq 0$, *where* h_1, *defined and differentiable on* $[1, 0]$, *satisfies* $h_1(0) = 0$ *and* $h_1(t) > 0$, $l - h_1'(t) \geqq 0$, $-1 \leqq t < 0$;

$$f_{y'}(t, y, u_R'(t), \varepsilon) \leqq h_2(t), \qquad |y - u_R(t)| \leqq d(t), \qquad 0 \leqq t \leqq 1,$$

where h_2, *defined and differentiable on* $[0, 1]$, *satisfies* $h_2(0) = 0$ *and* $h_2(t) < 0$, $l - h_2'(t) \geqq 0$, $0 < t \leqq 1$;

6)

$$f(t, u_L(t), u_L'(t), \varepsilon) = O(\varepsilon), \qquad -1 \leqq t \leqq 0;$$
$$f(t, u_R(t), u_R'(t), \varepsilon) = O(\varepsilon), \qquad 0 \leqq t \leqq 1.$$

Then for each $\varepsilon > 0$ *sufficiently small,* $\varepsilon \leqq \varepsilon_0$, *the problem* (3.1), (3.2) *has a solution* $y = y(t, \varepsilon)$. *Moreover, the following estimates hold*:

(i) *if* $u_L(0) > u_R(0)$,

$$u_L(t) - (u_L(0) - u_R(0)) \exp\left[-\varepsilon^{-1} \int_t^0 h_1(s)\, ds\right] - \varepsilon^{1/2} \bar{\gamma} \leqq y(t, \varepsilon) \leqq u_L(t) + \varepsilon^{1/2} \bar{\gamma}, \qquad -1 \leqq t \leqq 0;$$

$$u_R(t) - \varepsilon^{1/2} \bar{\gamma} \leqq y(t, \varepsilon) \leqq u_R(t) + (u_L(0) - u_R(0)) \exp\left[\varepsilon^{-1} \int_0^t h_2(s)\, ds\right] + \varepsilon^{1/2} \bar{\gamma}, \qquad 0 \leqq t \leqq 1;$$

(ii) *if* $u_L(0) < u_R(0)$,

$$u_L(t) - \varepsilon^{1/2} \bar{\gamma} \leqq y(t, \varepsilon) \leqq u_L(t) + (u_R(0) - u_L(0)) \exp\left[-\varepsilon^{-1} \int_t^0 h_1(s)\, ds\right] + \varepsilon^{1/2} \bar{\gamma}, \qquad -1 \leqq t \leqq 0;$$

[1] Here and throughout the paper the symbol $E(t_1, t_2)$, $-1 < t_1 < t_2 < 1$, denotes the subregion of $E(-1, 1)$ in (t, y, y', ε)-space obtained by restricting t to the subinterval $[t_1, t_2]$.

$$u_R(t)-(u_R(0)-u_L(0))\exp\left[\varepsilon^{-1}\int_0^t h_2(s)\,ds\right]-\varepsilon^{1/2}\bar{\gamma}\leqq y(t,\varepsilon)$$
$$\leqq u_R(t)+\varepsilon^{1/2}\bar{\gamma},\qquad 0\leqq t\leqq 1.$$

Here $\bar{\gamma}>0$ is a computable constant independent of ε. In addition, if $u_L'(0)=u_R'(0)$, the terms of order $O(\varepsilon^{1/2})$ in (i) *and* (ii) *can be replaced by terms of order $O(\varepsilon)$.*

Proof. It is sufficient to prove the result in the case that $u_L(0)>u_R(0)$ since the reflected case $u_L(0)<u_R(0)$ may be reduced to this one by making the change of dependent variable $y\to -y$. Define then for ε in $(0,\varepsilon_0]$

$$\alpha(t,\varepsilon)=\begin{cases}u_L(t)-H_1(t,\varepsilon)-(\varepsilon l^{-1})^{1/2}(u_L'(0)-u_R'(0))S_+(t,\varepsilon)-\varepsilon\gamma l^{-1} & -1\leqq t\leqq 0,\\ u_R(t)-(\varepsilon l^{-1})^{1/2}(u_L'(0)-u_R'(0))-\varepsilon\gamma l^{-1}, & 0\leqq t\leqq 1,\end{cases}$$

and

$$\beta(t,\varepsilon)=\begin{cases}u_L(t)+\varepsilon\gamma l^{-1}, & -1\leqq t\leqq 0,\\ u_R(t)+H_2(t,\varepsilon)+\varepsilon\gamma l^{-1}, & 0\leqq t\leqq 1,\end{cases}$$

if $u_L'(0)\geqq u_R'(0)$; while if $u_L'(0)<u_R'(0)$, define

$$\alpha(t,\varepsilon)=\begin{cases}u_L(t)-H_1(t,\varepsilon)-\varepsilon\gamma l^{-1}, & -1\leqq t\leqq 0,\\ u_R(t)-\varepsilon\gamma l^{-1}, & 0\leqq t\leqq 1,\end{cases}$$

and

$$\beta(t,\varepsilon)=\begin{cases}u_L(t)+(\varepsilon l^{-1})^{1/2}(u_R'(0)-u_L'(0))+\varepsilon\gamma l^{-1}, & -1\leqq t\leqq 0,\\ u_R(t)+H_2(t,\varepsilon)+(\varepsilon l^{-1})^{1/2}(u_R'(0)-u_L'(0))S_-(t,\varepsilon)+\varepsilon\gamma l^{-1}, & 0\leqq t\leqq 1.\end{cases}$$

Here $H_1(t,\varepsilon)=(u_L(0)-u_R(0))\exp[-\varepsilon^{-1}\int_t^0 h_1(s)\,ds]$, $H_2(t,\varepsilon)=(u_L(0)-u_R(0))\cdot\exp[\varepsilon^{-1}\int_0^t h_2(s)\,ds]$, and $S_\pm(t,\varepsilon)=\exp[\pm(\varepsilon^{-1}l)^{1/2}t]$; also, $\gamma>0$ is a constant to be determined.

The inequalities $\alpha\leqq\beta$, $\alpha(-1,\varepsilon)\leqq A\leqq\beta(-1,\varepsilon)$ and $\alpha(1,\varepsilon)\leqq B\leqq\beta(1,\varepsilon)$ are obvious (recall that $u_L(-1)=A$ and $u_R(1)=B$). In addition, $D_l\alpha(0)\leqq D_r\alpha(0)$ and $D_l\beta(0)\geqq D_r\beta(0)$. We consider now just the case $u_L'(0)\geqq u_R'(0)$ since the other one is handled similarly. It remains to verify that α satisfies the correct differential inequality on $[-1,0)$ and $(0,1]$; the verification for β is analogous and is omitted.

By noting that $f(t,\alpha,\alpha',\varepsilon)$ may be written as

$$f(t,\alpha,\alpha',\varepsilon)=f(t,u,u',\varepsilon)+\{f(t,\alpha,u',\varepsilon)-f(t,u,u',\varepsilon)\}$$
$$+\{f(t,\alpha,\alpha',\varepsilon)-f(t,\alpha,u',\varepsilon)\}$$

and expanding the bracketed terms by Taylor's theorem, we obtain the decomposition

$$f(t,\alpha,\alpha',\varepsilon)=f(t,u,u',\varepsilon)+f_y(t,u+\theta_1(\alpha-u),u',\varepsilon)(\alpha-u)$$
$$+f_{y'}(t,\alpha,u',\varepsilon)(\alpha'-u')$$
$$+\tfrac{1}{2}f_{y'y'}(t,\alpha,u'+\theta_2(\alpha'-u'),\varepsilon)(\alpha'-u')^2,$$

where $0<\theta_1, \theta_2<1$. Thus on $[-1, 0)$, differentiating α and substituting into $\varepsilon\alpha''-f(t, \alpha, \alpha', \varepsilon)$ via this decomposition, we obtain

$$\begin{aligned}\varepsilon\alpha''-f(t,\alpha,\alpha',\varepsilon)=&\ \varepsilon u_L''-(h_1'+\varepsilon^{-1}h_1^2)H_1(t,\varepsilon)-(\varepsilon l)^{1/2}(u_L'(0)-u_R'(0))S_+(t,\varepsilon)\\&-f(t,u_L,u_L',\varepsilon)+f_y(t,\eta_1,u_L',\varepsilon)\\&\cdot\{H_1(t,\varepsilon)+(\varepsilon l^{-1})^{1/2}(u_L'(0)-u_R'(0))S_+(t,\varepsilon)+\varepsilon\gamma l^{-1}\}\\&+f_{y'}(t,\alpha,u_L',\varepsilon)\{\varepsilon^{-1}h_1H_1(t,\varepsilon)+(u_L'(0)-u_R'(0))S_+(t,\varepsilon)\}\\&-\tfrac{1}{2}f_{y'y'}(t,\alpha,\eta_2',\varepsilon)\{\varepsilon^{-1}h_1H_1(t,\varepsilon)+(u_L'(0)-u_R'(0))S_+(t,\varepsilon)\}^2,\end{aligned}$$

where $\eta_1=u_L+\theta_1(\alpha-u_L)$ and $\eta_2'=u_L'+\theta_2(\alpha'-u_L')$. Finally, making use of the assumptions of the theorem (in particular, $f_{y'}(t, \alpha, u_L', \varepsilon)\geqq h_1(t)$ and $-f_{y'y'}(t, \alpha, \eta_2', \varepsilon)\geqq 0$) and taking ε sufficiently small, we have the desired inequality $\varepsilon\alpha''-f(t, \alpha, \alpha', \varepsilon)\geqq 0$, provided $\gamma\geqq\max\{|u_L''(t)|, \varepsilon^{-1}|f(t, u_L(t), u_L'(t), \varepsilon)|, 1\}$. On the interval $(0, 1]$, we have trivially

$$\begin{aligned}\varepsilon\alpha''-f(t,\alpha,\alpha',\varepsilon)&=\varepsilon u_L''-f(t,u_L,u_L',\varepsilon)+f_y(t,\eta_1,u_L',\varepsilon)\\&\qquad\cdot\{(\varepsilon l^{-1})^{1/2}(u_L'(0)-u_R'(0))+\varepsilon\gamma l^{-1}\}\\&\geqq-\varepsilon|u_L''|-\varepsilon|\varepsilon^{-1}f(t,u_L,u_L',\varepsilon)|+\varepsilon\gamma\\&\geqq 0,\quad\text{by our choice of }\gamma.\end{aligned}$$

Consequently we may apply Theorem 2.1 to conclude the proof of Theorem 3.1.

Remark 3.1. The assumption that $f_y(t, \eta_1, \rho', \varepsilon)$, $\rho'=u_L'$ or u_R', is positively bounded away from zero amounts to a stability assumption on the roots u_L, u_R. Since in a neighborhood of $t=0$, these functions annihilate $f_{y'}$, stability is determined by the (linearized) coefficient of y, i.e., by f_y. However it may happen that u_L and u_R also annihilate f_y; then it is reasonable to measure the stability of these roots by means of f_{yy}, f_{yyy}, etc. Such considerations are related to stability requirements which are imposed on reduced roots of the problem $\varepsilon^2y''=h(t, y)$ (see, e.g., [1] and [8]). As an example of such a condition, assumption 4) may be replaced by 4') $\partial_y^j f(t, u_L, u_L', \varepsilon)=O(\varepsilon)$, $-1\leqq t\leqq 0$, $1\leqq j\leqq 2q$;

$$\begin{aligned}\partial_y^{2q+1}f(t,\eta_1,u_L',\varepsilon)&\geqq l>0, &&-1\leqq t\leqq 0;\\ \partial_y^j f(t,u_R,u_R',\varepsilon)&=O(\varepsilon), &&0\leqq t\leqq 1,\quad 1\leqq j\leqq 2q;\\ \partial_y^{2q+1}f(t,\eta_1,u_R',\varepsilon)&\geqq l>0, &&0\leqq t\leqq 1.\end{aligned}$$

The error term of order $O(\varepsilon)$ in the definition of α and β is then replaced by a term of order $O(\varepsilon^{(2q+1)^{-1}})$. This is proved most easily by expanding the term $f(t, \sigma, \rho', \varepsilon)-f(t, \rho, \rho', \varepsilon)$ to $(2q+1)$-terms by Taylor's theorem; here $\sigma=\alpha$ or β, $\rho=u_L$ or u_R.

Remark 3.2. Assumption 5) that $f_{y'}$ is bounded by a certain function of t is only required for t in $(-\delta, \delta)$. For simplicity of exposition, however, we assumed that $f_{y'}$ was globally bounded by h_1 or h_2. It is clearly enough to assume additionally that $f_{y'}(t, y, u_L', \varepsilon)\geqq 0$, $|y-u_L(t)|\leqq d(t)$, $-1\leqq t\leqq -\delta$, and $f_{y'}(t, y, u_R', \varepsilon)\leqq 0$, $|y-u_R(t)|\leqq d(t)$, $\delta\leqq t\leqq 1$.

Remark 3.3. The proof of the theorem reveals that if the partials $\partial_y^j f$ only satisfy $\partial_y^j f(t, \rho, \rho', \varepsilon)\geqq 0$, $\rho=u_L$ or u_R (e.g., if $f_y\equiv 0$), then a result similar to that of

Theorem 3.1 holds, provided $u_L'' = u_R'' \equiv 0$ and $u_L'(0) = u_R'(0)$. In addition, estimates (i) and (ii) are valid with the terms of order $O(\varepsilon^{1/2})$ (or $O(\varepsilon)$) absent.

Remark 3.4. Two examples of functions h_1, h_2 satisfying assumption 5) are:

(a)

$$h_1(t) = -k_1 t^{2n+1}, \qquad -1 \leqq t \leqq 0,$$
$$h_2(t) = -k_2 t^{2n+1}, \qquad 0 \leqq t \leqq 1, \quad n = 0, 1, 2, \cdots;$$

(b)

$$h_1(t) = k_1 t^{2n}, \qquad -1 \leqq t \leqq 0,$$
$$h_2(t) = -k_2 t^{2n}, \qquad 0 \leqq t \leqq 1, \quad n = 1, 2, \cdots.$$

Remark 3.5. If the roots u_L, u_R of the reduced equation satisfy both the equations $u_L(0) = u_R(0)$ and $u_L'(0) = u_R'(0)$, then under even weaker assumptions than those of the theorem, a solution $y = y(t, \varepsilon)$ of (3.1), (3.2) exists and satisfies $y(t, \varepsilon)(y'(t, \varepsilon)) \to u_L(t)(u_L'(t))$, $-1 \leqq t \leqq 0$, and $y(t, \varepsilon)(y'(t, \varepsilon)) \to u_R(t)(u'_R(t))$, $0 \leqq t \leqq 1$. That is, neither the solution nor its derivative exhibits any nonuniform behavior in $[-1, 1]$. However, it frequently happens in the case of generalized turning point problems and elsewhere that the roots u_L, u_R satisfy $u_L(0) = u_R(0)$ but $u_L'(0) \neq u_R'(0)$. Such behavior is closely related to the theorem of Haber and Levinson [5] and it will be examined in that context in the next section.

One of our principal assumptions in Theorem 3.1 was that $f_{y'}$ behaved like a certain function of t in a δ-neighborhood of $t = 0$. It is also of interest to consider briefly the case in which $f_{y'}$ behaves, for example, like a function of the solution itself, i.e., $f_{y'} = g(y)$. Look then at the following special case of (3.1), (3.2)

$$\varepsilon y'' = p(t, y)(y')^2 + h(t, y, y'), \qquad -1 < t < 1, \tag{3.4}$$

$$y(-1, \varepsilon) = A, \quad y(1, \varepsilon) = B, \quad A \neq B, \tag{3.5}$$

in which $tp(t, y) \geqq 0$, $-1 \leqq t \leqq 1$, for all y. Suppose that the associated boundary value problem

$$\varepsilon z'' = h(t, z, z'), \qquad -1 < t < 1, \tag{3.6}$$

$$z(-1, \varepsilon) = A, \quad z(1, \varepsilon) = B, \quad A > B, \tag{3.7}$$

has a solution $z = z(t, \varepsilon)$ with a shock layer structure at $t = 0$, i.e., $z(t, \varepsilon) \to A$, $-1 \leqq t < 0$, $z(t, \varepsilon) \to B$, $0 < t \leqq 1$, as $\varepsilon \to 0^+$. Then if $z'(0^\pm, \varepsilon)$ is nonpositive, it turns out that for the range of $\varepsilon > 0$ for which $z(t, \varepsilon)$ exists, a solution $y = y(t, \varepsilon)$ of (3.4), (3.5) exists and satisfies $y(t, \varepsilon) \to A$, $-1 \leqq t < 0$, $y(t, \varepsilon) \to B$, $0 < t \leqq 1$, as $\varepsilon \to 0^+$. Thus y possesses the same shock layer structure as the solution z of the "simpler" problem (3.6), (3.7). This result is easily proved by defining the functions

$$\alpha(t, \varepsilon) = \begin{cases} z_1(t, \varepsilon), & -1 \leqq t \leqq 0, \\ B, & 0 \leqq t \leqq 1, \end{cases} \qquad \beta(t, \varepsilon) = \begin{cases} A & -1 \leqq t \leqq 0, \\ z_2(t, \varepsilon), & 0 \leqq t \leqq 1, \end{cases}$$

where z_1 is a solution of (3.6) satisfying $z_1(-1, \varepsilon) = A$, $z_1(0, \varepsilon) = B$, $z_1'(0^-, \varepsilon) \leqq 0$ and z_2 is a solution of (3.6) satisfying $z_2(0, \varepsilon) = A$, $z_2(1, \varepsilon) = B$, $z_2'(0^+, \varepsilon) \leqq 0$. The analogous result in the case that $tp(t, y) \leqq 0$, $-1 \leqq t \leqq 1$, follows similarly if $z'_{1,2}(0^\mp, \varepsilon)$ is nonnegative and $A < B$.

The final transition layer phenomenon we study in this section involves the case in which the reduced equation $f(t, u, u', 0) = 0$ has two singular solutions u_1, u_2 which exist at least on $[-1, 0]$, $[0, 1]$ respectively. The term "singular" in this context simply means that $f_{y'}(t, u_i(t), u_i'(t), 0) \equiv 0$, $i = 1, 2$. Since u_1, u_2 annihilate $f_{y'}$ on their respective intervals of existence, their stability is determined by the partial derivatives $\partial_y^j f$, for a certain range of $j \geqq 1$. In addition, singular reduced roots cannot in general be made to satisfy either of the supplementary conditions of the original problem. Consequently, under hypotheses to be given below, a solution of the full problem (3.1), (3.2) can possess, in addition to shock layer behavior at $t = 0$, boundary layer behavior at both endpoints. Our earlier remarks about the relationship between the convexity properties of solutions inside a shock layer apply equally to the case of a boundary layer. For instance, if $f_{y'y'}$ is positive near $t = -1$, then a solution y of (3.1), (3.2) is convex in any boundary layer at $t = -1$, i.e., $y'' > 0$. The singular solution u_1 must then satisfy an inequality of the form $u_1(-1) < A$, if it is to represent y to the edge of the boundary layer. We give these intuitive ideas a precise expression in the next theorem, which actually applies to a broader class of reduced roots than the singular ones discussed above. It is stated for the case in which $f_{y'y'}$ is negative near $t = -1$ and positive near $t = 1$.

THEOREM 3.2. *Assume*

1) *the reduced equation* $f(t, u, u', 0) = 0$ *has two solutions* $u_1 = u_1(t)$ *and* $u_2 = u_2(t)$, *defined and of class* $C^{(2)}$ *on* $[-1, 0]$, $[0, 1]$, *respectively, with* $u_1(0) \neq u_2(0)$ *and* $u_1(-1) > A$, $u_2(1) < B$;

2) *the function f is continuous in* (t, y, y', ε) *and of class* $C^{(2)}$ *with respect to* y, y' *in* $\mathscr{D}(-1, 1)$: $-1 \leqq t \leqq 1$, $|y - u_1(t)| \leqq d_1(t)$, $-1 \leqq t \leqq 0$, $d_1(t) \geqq u_1(-1) - A$, $-1 \leqq t < -1 + \delta$, $d_1(t) = O(\varepsilon)$, $-1 + \delta \leqq t \leqq -\delta$, $d_1(t) \geqq |u_1(0) - u_2(0)|$, $-\delta < t \leqq 0$, $|y - u_2(t)| \leqq d_2(t)$, $0 \leqq t \leqq 1$, $d_2(t) \geqq |u_1(0) - u_2(0)|$, $0 \leqq t < \delta$, $d_2(t) = O(\varepsilon)$, $\delta \leqq t \leqq 1 - \delta$, $d_2(t) \geqq B - u_2(1)$, $1 - \delta < t \leqq 1$, $|y'| < \infty$, $0 \leqq \varepsilon \leqq \varepsilon_0$, $\varepsilon_0 > 0$; *and in* $\mathscr{D}(-1, 1)$, $f = O(|y'|^2)$, *as* $|y'| \to \infty$; *here* $\delta > 0$ *is a small constant independent of* ε;

3) *for the same constant* $\delta > 0$ *in* 2), $(u_1(0) - u_2(0))f_{y'y'} \leqq 0$ *in* $\mathscr{D}(-\delta, 0)$ *and* $(u_1(0) - u_2(0))f_{y'y'} \geqq 0$ *in* $\mathscr{D}(0, \delta)$;

4) *there is a positive constant p such that* $f_{y'y'} \leqq -p < 0$ *in* $\mathscr{D}(-1, -1 + \delta)$ *and* $f_{y'y'} \geqq p > 0$ *in* $\mathscr{D}(1 - \delta, 1)$;

5) *there is a positive constant l such that*

$$f_y(t, y, u_1'(t), \varepsilon) \geqq l > 0, \qquad |y - u_1(t)| \leqq d_1(t), \qquad -1 \leqq t \leqq 0;$$

$$f_y(t, y, u_2'(t), \varepsilon) \geqq l > 0, \qquad |y - u_2(t)| \leqq d_2(t), \qquad 0 \leqq t \leqq 1;$$

6)

$$f_{y'}(t, y, u_1'(t), \varepsilon) \geqq 0, \qquad |y - u_1(t)| \leqq d_1(t), \qquad -\delta \leqq t \leqq 0,$$

$$f_{y'}(t, y, u_2'(t), \varepsilon) \leqq 0, \qquad |y - u_2(t)| \leqq d_2(t), \qquad 0 \leqq t \leqq \delta;$$

and

$$f_{y'}(t, u_1(t), u_1'(t), \varepsilon) \leqq 0, \qquad -1 \leqq t \leqq -1 + \delta,$$

$$f_{y'}(t, u_2(t), u_2'(t), \varepsilon) \geqq 0, \qquad 1 - \delta \leqq t \leqq 1;$$

7)

$$f(t, u_1(t), u_1'(t), \varepsilon) = O(\varepsilon), \qquad -1 \leqq t \leqq 0,$$

$$f(t, u_2(t), u_2'(t), \varepsilon) = O(\varepsilon), \qquad 0 \leqq t \leqq 1.$$

Then for each $\varepsilon>0$ *sufficiently small,* $\varepsilon\leqq\varepsilon_0$, *the problem* (3.1), (3.2) *has a solution* $y=y(t,\varepsilon)$. *Moreover, if* $u_1(0)>u_2(0)$,

$$u_1(t)-(u_1(-1)-A)\exp[-(\varepsilon^{-1}l)^{1/2}(1+t)]-(u_1(0)-u_2(0))\exp[(\varepsilon^{-1}l)^{1/2}t]$$
$$-\varepsilon\bar{\gamma}\leqq y(t,\varepsilon)\leqq u_1(t)+\varepsilon\bar{\gamma},\qquad -1\leqq t\leqq 0,$$
$$u_2(t)-\varepsilon\bar{\gamma}\leqq y(t,\varepsilon)\leqq u_2(t)+(B-u_2(1))\exp[-(\varepsilon^{-1}l)^{1/2}(1-t)]$$
$$+(u_1(0)-u_2(0))\exp[-(\varepsilon^{-1}l)^{1/2}t]+\varepsilon\bar{\gamma},\qquad 0\leqq t\leqq 1;$$

while if $u_1(0)<u_2(0)$,

$$u_1(t)-(u_1(-1)-A)\exp[-(\varepsilon^{-1}l)^{1/2}(1+t)]-\varepsilon\bar{\gamma}\leqq y(t,\varepsilon)$$
$$\leqq u_1(t)+(u_2(0)-u_1(0))\exp[(\varepsilon^{-1}l)^{1/2}t]+\varepsilon\bar{\gamma},\qquad -1\leqq t\leqq 0,$$
$$u_2(t)-(u_2(0)-u_1(0))\exp[-(\varepsilon^{-1}l)^{1/2}t]-\varepsilon\bar{\gamma}\leqq y(t,\varepsilon)$$
$$\leqq u_2(t)+(B-u_2(1))\exp[-(\varepsilon^{-1}l)^{1/2}(1-t)]+\varepsilon\bar{\gamma},\qquad 0\leqq t\leqq 1.$$

Proof. In the case that $u_1(0)>u_2(0)$, the proof follows as that of Theorem 3.1 by defining the functions, for $0<\varepsilon\leqq\varepsilon_0$:

$$\alpha(t,\varepsilon)=\begin{cases} u_1(t)-(u_1(-1)-A)\exp[-(\varepsilon^{-1}l)^{1/2}(1+t)] \\ \quad -(u_1(0)-u_2(0))\exp[(\varepsilon^{-1}l)^{1/2}t]-\varepsilon\gamma l^{-1} \\ \quad +(u_1(-1)-A)\exp[-(\varepsilon^{-1}l)^{1/2}], & -1\leqq t\leqq 0, \\ u_2(t)-\varepsilon\gamma l^{-1}, & 0\leqq t\leqq 1, \end{cases}$$

$$\beta(t,\varepsilon)=\begin{cases} u_1(t)+\varepsilon\gamma l^{-1}, & -1\leqq t\leqq 0, \\ u_2(t)+(B-u_2(1))\exp[-(\varepsilon^{-1}l)^{1/2}(1-t)] \\ \quad +(u_1(0)-u_2(0))\exp[-(\varepsilon^{-1}l)^{1/2}t]+\varepsilon\gamma l^{-1} \\ \quad -(B-u_2(1))\exp[-(\varepsilon^{-1}l)^{1/2}], & 0\leqq t\leqq 1. \end{cases}$$

The reflected case $u_1(0)<u_2(0)$ is handled similarly. Note that for ε sufficiently small, $\alpha\leqq\beta$ and

$$D_l\alpha(0)=u_1'(0^-)+l^{1/2}\varepsilon^{-1/2}(u_1(-1)-A)\exp[-(\varepsilon^{-1}l)^{1/2}]$$
$$-l^{1/2}\varepsilon^{-1/2}(u_1(0)-u_2(0))<D_r\alpha(0)=u_2'(0^+);$$

similarly,

$$D_l\beta(0)>D_r\beta(0).$$

Remark 3.6. Analogous results are valid if instead of requiring $u_1(-1)>A$ and $u_2(1)<B$, we assume that $u_1(-1)<A$ and $u_2(1)<B$, or $u_1(-1)<A$ and $u_2(1)>B$, or $u_1(-1)>A$ and $u_2(1)>B$. In each of these cases, we must assume additionally that $f_{y'y'}>0$ in $\mathscr{D}(-1,1+\delta)$ and $f_{y'y'}>0$ in $\mathscr{D}(1-\delta,1)$, or $f_{y'y'}>0$ in $\mathscr{D}(-1,-1+\delta)$ and $f_{y'y'}<0$ in $\mathscr{D}(1-\delta,1)$, or $f_{y'y'}<0$ in $\mathscr{D}(-1,-1+\delta)$ and $f_{y'y'}<0$ in $\mathscr{D}(1-\delta,1)$, respectively.

Remark 3.7. The stability assumption, assumption 4) of Theorem 3.2, may be modified as in [9] to include the case of roots u_1, u_2 which annihilate a succession of the functions $\partial_y^j f$, for a range of $j \geqq 1$. It is then necessary to use algebraic boundary and interior layer functions (cf. [8] and [9]) in place of exponential functions.

Remark 3.8. It is clear that the theory presented in Theorems 3.1 and 3.2 may be combined to treat the case of a "regular-singular" shock layer. Namely the reduced equation has a regular solution, say $u = u_L(t)$, on $[-1, 0]$ with $u_L(-1) = A$ and $f_{y'}[u_L(t)] \geqq h_1(t)$, and a stable singular solution $u = u_S(t)$ on $[0, 1]$ such that $u_L(0) \neq u_S(0)$. Then under the obvious assumptions, the full problem has a solution $y = y(t, \varepsilon)$, for small enough ε, which exhibits shock layer behavior at $t = 0$ and possibly boundary layer behavior at $t = 1$, i.e., $y(t, \varepsilon) \to u_L(t)$, $-1 \leqq t < 0$, $y(t, \varepsilon) \to u_S(t)$, $0 < t < 1$, as $\varepsilon \to 0^+$.

We conclude this section with five examples which illustrate the theory given above.

Example 3.1. Consider the problem

$$\varepsilon y'' = t(y')^2 - ty', \qquad -1 < t < 1,$$
$$y(-1, \varepsilon) = A, \quad y(1, \varepsilon) = B, \quad A > B.$$

It is easy to verify that $u_L \equiv A$ and $u_R \equiv B$ satisfy the hypotheses of Theorem 3.1 with $h_1(t) = h_2(t) = -t$ (cf. Remark 3.3). For each $\varepsilon > 0$, a solution $y = y(t, \varepsilon)$ exists and satisfies

$$A - (A - B)\exp[-(2\varepsilon)^{-1}t^2] \leqq y(t, \varepsilon) \leqq A, \qquad -1 \leqq t \leqq 0,$$
$$B \leqq y(t, \varepsilon) \leqq B + (A - B)\exp[-(2\varepsilon)^{-1}t^2], \qquad 0 \leqq t \leqq 1.$$

Example 3.2. For the problem

$$\varepsilon y'' = f(t, y, y'') = \begin{cases} t(y')^2 + tyy', & -1 \leqq t \leqq 0, \\ t(y')^2 - tyy', & 0 \leqq t \leqq 1, \end{cases}$$
$$y(-1, \varepsilon) = A; \quad y(1, \varepsilon) = B, \quad Ae^{-1} > B > 0,$$

the reduced equation $f(t, u, u') = 0$ has the roots $\bar{u}_L \equiv A$, $u_L(t) = A\exp[-(1+t)]$, for t in $[-1, 0]$, and $\bar{u}_R \equiv B$, $u_R(t) = B\exp[t - 1]$, for t in $[0, 1]$. Of these, u_L and $\bar{u}_R$ form a stable pair since $f_{y'}(t, \eta_1, u_L'(t)) \geqq -Ae^{-1}t$, $-1 \leqq t \leqq 0$, for $\eta_1 \leqq A\exp[-(1+t)]$, and $f_{y'}(t, \eta_2, \bar{u}_R'(t)) \leqq -Bt$, $0 \leqq t \leqq 1$, for $\eta_2 \geqq B$. Setting $h_1(t) = -Ae^{-1}t$ and $h_2(t) = -Bt$ we argue as in the proof of Theorem 3.1 (cf. Remark 3.3) to deduce that this problem has a solution $y = y(t, \varepsilon)$, for ε sufficiently small, which satisfies

$$A\exp[-(1+t)] - (Ae^{-1} - B)\exp[-Ae^{-1}(2\varepsilon)^{-1}t^2] - \varepsilon^{1/2}\bar{\gamma} \leqq y(t, \varepsilon)$$
$$\leqq A\exp[-(1+t)], \qquad -1 \leqq t \leqq 0,$$
$$B - \varepsilon^{1/2}\bar{\gamma} \leqq y(t, \varepsilon) \leqq B + (Ae^{-1} - B)\exp[-B(2\varepsilon)^{-1}t^2] + \varepsilon^{1/2}\bar{\gamma}, \qquad 0 \leqq t \leqq 1.$$

Theorem 3.1 is not directly applicable here since $f_y(t, y, u_L'(t)) = -tA\exp[-(1+t)] \geqq 0$, $-1 \leqq t \leqq 0$, and $f_y(t, y, \bar{u}_R'(t)) \equiv 0$, $0 \leqq t \leqq 1$.

Example 3.3. Consider next

$$\varepsilon y''=t(y')^2+yy', \qquad -1<t<1,$$
$$y(-1,\varepsilon)=A, \quad y(1,\varepsilon)=B, \quad A=-B>0.$$

The associated problem $\varepsilon z''=zz'$, $-1<t<1$, $z(-1,\varepsilon)=A, z(1,\varepsilon)=B$, can be solved explicitly to yield a solution satisfying $z(t,\varepsilon)\to A$, $-1\leqq t<0$, $z(t,\varepsilon)\to B$, $0<t\leqq 1$. We can then apply the theory discussed above to show that the problem for y has a solution $y(t,\varepsilon)$ satisfying $y(t,\varepsilon)\to A, -1\leqq t<0, y(t,\varepsilon)\to B, 0<t\leqq 1$.

Example 3.4. This example is actually a whole family of problems. Consider

$$\varepsilon y''=t(y')^2+h(t,y), \qquad -1<t<1,$$
$$y(-1,\varepsilon)=A, \qquad y(1,\varepsilon)=B,$$

where $h(t,u)=0$ has two constant roots u_1, u_2 such that $h_y(t,u_i)>0$, $i=1,2$. Then if the quantities $|u_1-u_2|$, $|A-u_1|$ and $|B-u_2|$ are sufficiently small, we can apply the reasoning behind Theorem 3.2 to deduce the existence of a solution $y=y(t,\varepsilon)$ such that $y(t,\varepsilon)\to u_1, -1<t<0, y(t,\varepsilon)\to u_2, 0<t<1$.

Example 3.5. Our final example gives an illustration of a "regular-singular" shock layer which was briefly described in Remark 3.8. Consider the problem

$$\varepsilon y''=y-t(y')^2, \qquad -1<t<1,$$
$$y(-1,\varepsilon)=A, \quad y(1,\varepsilon)=B, \quad A<-1, \quad B\leqq 0.$$

The reduced pair of interest is $u_L(t)=-((-A)^{1/2}-1+(-t)^{1/2})^2, -1\leqq t\leqq 0$, and $u_S\equiv 0$, $0\leqq t\leqq 1$. Since $f_{y'}(t,y,u_L'(t))\geqq h_1(t)=2((-A^{1/2}-1)(-t)^{1/2}$, $-1\leqq t\leqq 0$, and $f_y\equiv 1$, the roots u_L and u_S are stable. It is then possible to argue as above and show that this problem has a unique solution $y=y(t,\varepsilon)$ satisfying

$$y(t,\varepsilon)=u_L(t)+O\Big(\exp\Big[-\varepsilon^{-1}\int_t^0 h_1(s)\,ds\Big]\Big)+O(\varepsilon^{1/2}), \qquad -1\leqq t\leqq 0,$$
$$y(t,\varepsilon)=O(\exp[-\varepsilon^{-1/2}t])+O(|B|\exp[-\varepsilon^{-1/2}(1-t)])+O(\varepsilon^{1/2}), \qquad 0\leqq t\leqq 1.$$

The argument is not as straightforward as previous ones since u_L'' is unbounded at $t=0$, but this offers no real difficulty.

4. Modified Haber–Levinson crossings. We consider in this section phenomena which were first studied in the classic paper of Haber and Levinson [5]. (See also [16, Chap. 2], [14] and [10] for more recent discussions.) The situation here is the following. The reduced equation $f(t,u,u',0)=0$ is assumed to have two roots $u=u_L(t)$ and $u=u_R(t)$ which exist on $[-1,t_0]$ and $[t_0,1]$, respectively, $-1<t_0<1$, and satisfy $u_L(-1)=A$, $u_R(1)=B$. In addition, these roots are assumed to have an "angular" crossing at $t=t_0$, i.e., $u_L(t_0)=u_R(t_0)$ and $u_L'(t_0)\neq u_R'(t_0)$. Finally if u_L and u_R are stable in the sense that $f_{y'}(t,u_L(t),u_L'(t),0)\geqq k>0$ on $[-1,t_0]$ and $f_{y'}(t,u_R(t),u_R'(t),0)\leqq -k<0$ on $[t_0,1]$, and if

$$f(t_0,u_L(t_0),\omega,0)\begin{cases}>0, & u_L'(t_0)<\omega<u_R'(t_0),\\ <0, & u_R'(t_0)<\omega<u_L'(t_0),\end{cases}$$

Haber and Levinson showed that the problem

$$(4.1)\qquad \varepsilon y''=f(t, y, y', \varepsilon), \qquad -1<t<1,$$

$$(4.2)\qquad y(-1, \varepsilon)=A, \qquad y(1, \varepsilon)=B,$$

has a locally unique solution $y=y(t, \varepsilon)$ for small enough $\varepsilon>0$. In addition,

$$\lim_{\varepsilon\to 0^+} y(t, \varepsilon)=\begin{cases} u_L(t), & -1\leqq t\leqq t_0,\\ u_R(t), & t_0\leqq t\leqq 1;\end{cases} \qquad \lim_{\varepsilon\to 0^+} y'(t, \varepsilon)=\begin{cases} u_L'(t), & -1\leqq t\leqq t_0-\delta,\\ u_R'(t), & t_0+\delta\leqq t\leqq 1,\end{cases}$$

for $\delta>0$ a small constant independent of ε. This theorem of Haber and Levinson is of course valid without any form of growth condition on f with respect to y'.

For problems (4.1), (4.2) with generalized turning points whose right-hand sides satisfy $f=O(|y'|^2)$, as $|y'|\to\infty$, it frequently happens that there exist solutions which behave like those described by Haber and Levinson, despite the fact that f does not satisfy the original assumptions of [5]. In the next three theorems we investigate the nature of these modified assumptions. The proofs are omitted since this entire question is discussed more fully in [10]. However, with slight adjustment, the proof of Theorem 4.2 follows from that of Theorem 3.1.

THEOREM 4.1. *Assume*

1) *the reduced equation* $f(t, u, u', 0)=0$ *has two solutions* $u=u_L(t)$, $u=u_R(t)$ *which are defined and of class* $C^{(2)}$ *on* $[-1, t_0]$, $[t_0, 1]$, *respectively*, $t_0\neq 0$, *with* $u_L(-1)=A$, $u_R(1)=B$, $u_L(t_0)$, *and* $u_L'(t_0)\neq u_R'(t_0)$;

2) *the function* f *is continuous in* (t, y, y', ε) *and of class* $C^{(2)}$ *with respect to* y, y' *in* $\mathcal{R}: -1\leqq t\leqq 1$, $|y-u_L(t)|\leqq d_1$, $-1\leqq t\leqq t_0$, $|y-u_R(t)|\leqq d_2$, $t_0\leqq t\leqq 1$, $d_1, d_2=O(\varepsilon)$, $|y'|<\infty$, $0\leqq\varepsilon\leqq\varepsilon_0$, $\varepsilon_0>0$; *and in* $\mathcal{R}$, $f_{y'y'}=O(1)$, *as* $|y'|\to\infty$; $f_{y'y'}(0, \eta, \eta', \varepsilon)=0$, $(0, \eta, \eta', \varepsilon)$ *in* $\mathcal{R}$;

3) *there are positive constants* k_1, k_2, l *and* δ *such that*

$$f_{y'}(t, u_L(t), u_L'(t), 0)\geqq k_1>0, \qquad t_0-\delta\leqq t\leqq t_0,$$

$$f_{y'}(t, u_R(t), u_R'(t), 0)\leqq -k_2<0, \qquad t_0\leqq t\leqq t_0+\delta;$$

$$f_{y}(t, u_L(t), u_L'(t), 0)\geqq l>0, \qquad -1\leqq t\leqq t_0-\delta,$$

$$f_{y}(t, u_R(t), u_R'(t), 0)\geqq l>0, \qquad t_0+\delta\leqq t\leqq 1;$$

moreover,

$$f_{y'}(t, u_L(t), u_L'(t), \varepsilon)\geqq 0, \qquad -1\leqq t\leqq t_0-\delta,$$

$$f_{y'}(t, u_R(t), u_R'(t), \varepsilon)\leqq 0, \qquad t_0+\delta\leqq t\leqq 1;$$

4)

$$f(t_0, u_L(t_0), \omega, 0)\begin{cases} >0, & u_L'(t_0)<\omega<u_R'(t_0),\\ <0, & u_R'(t_0)<\omega<u_L'(t_0);\end{cases}$$

5)

$$f(t, u_L(t), u_L'(t), \varepsilon)=O(\varepsilon), \qquad -1\leqq t\leqq t_0,$$

$$f(t, u_R(t), u_R'(t), \varepsilon)=O(\varepsilon), \qquad t_0\leqq t\leqq 1.$$

Then for each $\varepsilon>0$ *sufficiently small*, $\varepsilon\leqq\varepsilon_0$, *the problem* (4.1), (4.2) *has a solution* $y=y(t, \varepsilon)$ *such that* $y(t, \varepsilon)=u_L(t)+O(\varepsilon)$, $-1\leqq t\leqq t_0$, *and* $y(t, \varepsilon)=u_R(t)+O(\varepsilon)$, $t_0\leqq t\leqq 1$.

The theorem says that the strict nonvanishing of the function $f_{y'}$ along the appropriate reduced path is required only in a small neighborhood of $t=t_0$ provided that f_y is strictly positive outside of this neighborhood. We remark that an angular crossing cannot occur at $t_0=0$ since $f_{y'y'}(0,\eta,\eta',\varepsilon)=0$ and $f_{y'}(t_0,\rho(t_0),\rho'(t_0),0)\neq 0$, $\rho=u_L$, u_R, i.e., if $u_L(0)=u_R(0)$, then $u_L'(0)=u_R'(0)$, as follows from an easy application of Taylor's theorem.

In the next theorem it is shown that a Haber–Levinson crossing can occur at a point t_0 at which $f_{y'y'}$ and $f_{y'}$ vanish simultaneoulsy. For convenience, we take $t_0=0$.

THEOREM 4.2. *Assume* 1), 2) *and* 5) *as in Theorem* 4.1 *with* $t_0=0$ *and* $d_1, d_2=O(\varepsilon^{1/2})$. Assume also

3)

$$f_{y'}(t,u_L(t),u_L'(t),\varepsilon)\geqq 0, \qquad -1\leqq t\leqq 0,$$

$$f_{y'}(t,u_R(t),u_R'(t),\varepsilon)\leqq 0, \qquad 0\leqq t\leqq 1;$$

4) *there is a constant* $l>0$ *such that*

$$f_y(t,u_L(t),u_L'(t),0)\geqq l>0, \qquad -1\leqq t\leqq 0,$$

$$f_y(t,u_R(t),u_R'(t),0)\geqq l>0, \qquad 0\leqq t\leqq 1.$$

Then for each $\varepsilon>0$ *sufficiently small,* $\varepsilon\leqq\varepsilon_0$, *the problem* (4.1), (4.2) *has a solution* $y=y(t,\varepsilon)$ *such that*

$$y(t,\varepsilon)=u_L(t)+O(\varepsilon^{1/2}), \qquad -1\leqq t\leqq 0,$$

$$y(t,\varepsilon)=u_R(t)+O(\varepsilon^{1/2}), \qquad 0\leqq t\leqq 1.$$

This theorem once again shows the interplay between the functions f_y and $f_{y'}$ which seems to typify linear and nonlinear turning point phenomena (cf. [7]). We note that the error term is of order $O(\varepsilon^{1/2})$ since we have only assumed that $f_{y'}$ is nonnegative (nonpositive) along $u_L(u_R)$.

We consider finally a case in which the reduced equation has two singular solutions which have an angular crossing at $t=0$. Such solutions in general satisfy neither boundary condition, so that in addition to Haber–Levinson behavior at $t=0$, the solution of the full problem exhibits boundary layer behavior at $t=\pm 1$.

THEOREM 4.3. *Assume*

1) *the reduced equation* $f(t,u,u',0)=0$ *has two singular solutions* $u=u_1(t)$, $u=u_2(t)$, *defined and of class* $C^{(2)}$ *on* $[-1,0]$, $[0,1]$, *respectively, with* $u_1(0)=u_2(0)$, $u_1'(0)\neq u_2'(0)$, *and* $u_1(-1)\leqq A$, $u_2(1)\geqq B$;

2) *the function* f *is continuous in* (t,y,y',ε) *and of class* $C^{(2)}$ *with respect to* y, y' *in* $\mathscr{D}$ $(-1,1)$: $-1\leqq t\leqq 1$, $|y-u_1(t)|\leqq d_1(t)$, $-1\leqq t\leqq 0$, $d_1(t)\geqq A-u_1(-1)$, $-1\leqq t<-1+\delta$, $d_1(t)=O(\varepsilon^{1/2})$, $-1+\delta\leqq t\leqq 0$, $|y-u_2(t)|\leqq d_2(t)$, $0\leqq t\leqq 1$, $d_2(t)=O(\varepsilon^{1/2})$, $0\leqq t\leqq 1-\delta$, $d_2(t)\geqq u_2(1)-B$, $1-\delta<t\leqq 1$, *for* $\delta>0$ *a small constant,* $|y'|<\infty$, $0\leqq\varepsilon\leqq\varepsilon_0$, $\varepsilon_0>0$; *in* $\mathscr{D}$ $(-1,1)$, $f_{y'y'}=O(1)$ *as* $|y'|\to\infty$, *and* $f_{y'y'}(0,\eta,\eta',\varepsilon)=0$, $(0,\eta,\eta',\varepsilon)$ *in* $\mathscr{D}$;

3) *there is a constant* $p>0$ *such that* $f_{y'y'}\geqq p>0$ *in* $\mathscr{D}(-1,-1+\delta)$ *and* $f_{y'y'}\leqq -p<0$ *in* $\mathscr{D}(1-\delta,1)$, *for* $\delta>0$ *as in* 2);

4) *there is a constant* $l > 0$ *such that*

$$f_y(t, y, u_1'(t), 0) \geqq l > 0, \qquad |y - u_1(t)| \leqq d_1(t), \qquad -1 \leqq t \leqq 0,$$
$$f_y(t, y, u_2'(t), 0) \geqq l > 0, \qquad |y - u_2(t)| \leqq d_2(t), \qquad 0 \leqq t \leqq 1;$$

5)

$$f(t, u_1(t), u_1'(t), \varepsilon) = O(\varepsilon), \qquad -1 \leqq t \leqq 0,$$
$$f(t, u_2(t), u_2'(t), \varepsilon) = O(\varepsilon), \qquad 0 \leqq t \leqq 1.$$

Then for each $\varepsilon > 0$ *sufficiently small,* $\varepsilon \leqq \varepsilon_0$, *the problem* (4.1), (4.2) *has a solution* $y = y(t, \varepsilon)$ *such that*

$$u_1(t) - \varepsilon^{1/2}\gamma \leqq y(t, \varepsilon) \leqq u_1(t) + (A - u_1(-1)) \exp[-(\varepsilon^{-1}l)^{1/2}(1+t)] + \varepsilon^{1/2}\gamma, \qquad -1 \leqq t \leqq 0,$$

$$u_2(t) - (u_2(1) - B) \exp[-(\varepsilon^{-1}l)^{1/2}(1-t)] - \varepsilon^{1/2}\gamma \leqq y(t, \varepsilon) \leqq u_2(t) + \varepsilon^{1/2}\gamma, \qquad 0 \leqq t \leqq 1.$$

Remark 4.1. The theorem of Haber and Levinson contains an estimate on the derivative $y'(t, \varepsilon)$ of the solution of (4.1), (4.2) (see also [16, Chap. 2] and [14]). In Theorems 4.1–4.3 we can also give a similar estimate for $y'(t, \varepsilon)$ based upon the estimates for $y - u_L$ and $y - u_R$.

Remark 4.2. The reduced roots were assumed to satisfy the condition that f_y was strictly positive in a certain domain including the reduced paths. It is frequently useful to consider the various cases in which such roots annihilate several of the functions $\partial_y^j f$. These more general results are discussed in [10]. With regard to such an extension of Theorem 4.3, see the discussion in [9, § 3].

Remark 4.3. We remark finally that the boundary layer behavior observed in Theorem 4.3 is only possible under the assumption that $f_{y'y'} = O(1)$. An analogous result is valid if instead of assuming $u_1(-1) \leqq A$, $u_2(1) \geqq B$, we assume: $u_1(-1) \leqq A$, $u_2(1) \leqq B$ and $f_{y'y'} > 0$ in $\mathscr{D}(-1, -1+\delta)$, $f_{y'y'} > 0$ in $\mathscr{D}(1-\delta, 1)$, or $u_1(-1) \geqq A$, $u_2(1) \leqq B$ and $f_{y'y'} < 0$ in $\mathscr{D}(-1, -1+\delta)$, $f_{y'y'} > 0$ in $\mathscr{D}(1-\delta, 1)$, or $u_1(-1) \geqq A$, $u_2(1) \geqq B$ and $f_{y'y'} < 0$ in $\mathscr{D}(-1, -1+\delta)$, $f_{y'y'} < 0$ in $\mathscr{D}(1-\delta, 1)$.

We conclude this section with three examples.

Example 4.1. Consider the problem

$$\varepsilon y'' = y^2 - t^2(y')^2, \qquad -1 < t < 1,$$
$$y(-1, \varepsilon) = A, \quad y(1, \varepsilon) = B, \quad -B < A < 0.$$

The stable solutions of the reduced equation for such A and B are $u_L(t) = -At^{-1}$, $-1 \leqq t < 0$, and $u_R(t) = Bt$, $-1 \leqq t \leqq 1$. Indeed, $f_{y'}[u_L(t)] = -2A > 0$ and $f_{y'}[u_R(t)] = -2Bt^2 < 0$, $t \neq 0$. Clearly, $u_L(t_0) = u_R(t_0)$, for $t_0 = -(-AB^{-1})^{1/2}$ in $(-1, 0)$, and $u_L'(t_0) = -B <'_R(t_0) = B$. Finally the crossing condition is satisfied, i.e., $(Bt_0)^2 - t_0^2\omega^2 = t_0^2(B^2 - \omega^2) > 0$, $|\omega| < B$. Thus we can apply Theorem 4.1 to conclude that this problem has a solution $y = y(t, \varepsilon)$ satisfying

$$u(t) - \varepsilon^{1/2}\gamma \leqq y(t, \varepsilon) \leqq u(t), \qquad -1 \leqq t \leqq 1,$$

where

$$u(t)=\begin{cases}-At^{-1}, & -1\leqq t\leqq t_0,\\ Bt, & t_0\leqq t\leqq 1.\end{cases}$$

Example 4.2. Consider now

$$\varepsilon y''=t(y')^2-yy', \qquad -1<t<1,$$

$$y(-1,\varepsilon)=A, \quad y(1,\varepsilon)=B, \quad A>0, \quad B<0, \quad A\neq -B.$$

The reduced equation has the stable pair of roots $u_L(t)=-At$ and $u_R(t)=Bt$, since $f_{y'}[u_L(t)]=-At\geqq 0, -1\leqq t\leqq 0$, and $f_{y'}[u_R(t)]=Bt\leqq 0, 0\leqq t\leqq 1$. Clearly $u_L(0)=u_R(0)$ and $u_L'(0)=-A\neq u_R'(0)=B$; moreover, $f_y[u_L(t)]=A>0$ and $f_y[u_R(t)]=-B>0$. We can then apply Theorem 4.2 to conclude that this problem has a solution $y=y(t,\varepsilon)$, for small enough $\varepsilon>0$, satisfying

$$u(t)\leqq y(t,\varepsilon)\leqq u(t)+\varepsilon^{1/2}\gamma, \qquad -1\leqq t\leqq 1, \quad \text{if } -A<B,$$

$$u(t)-\varepsilon^{1/2}\gamma\leqq y(t,\varepsilon)\leqq u(t), \qquad -1\leqq t\leqq 1, \quad \text{if } -A>B,$$

where

$$u(t)=\begin{cases}-At, & -1\leqq t\leqq 0,\\ Bt, & 0\leqq t\leqq 1.\end{cases}$$

Example 4.3. Consider finally the problem

$$\varepsilon y''=|t|(y')^2-2ty'+y, \qquad -1<t<1,$$

$$y(-1,\varepsilon)=A, \quad y(l,\varepsilon)=B, \quad A,B\geqq 1.$$

The reduced roots in this case are the singular ones $u_1(t)=-t, -1\leqq t\leqq 0$, and $u_2(t)=t, 0\leqq t\leqq 1$. This pair is clearly stable since $f_y\equiv 1$, and since $u_1(-1)\leqq A, u_2(1)\leqq B, f_{y'y'}\equiv 2|t|\geqq 0$, we can apply Theorem 4.3 to conclude that this problem has a unique solution $y=y(t,\varepsilon)$ for which

$$|t|\leqq y(t,\varepsilon)\leqq |t|+(A-1)\exp[-\varepsilon^{-1/2}(1+t)]$$
$$+(B-1)\exp[-\varepsilon^{-1/2}(1-t)]+\varepsilon^{1/2}\gamma, \qquad -1\leqq t\leqq 1.$$

5. Boundary layer phenomena. In this section we consider the case in which a single solution of the reduced equation generates a solution of the full problem (4.1), (4.2) that exhibits boundary layer behavior at $t=-1, t=1$ or both endpoints. Under the assumptions given below it turns out that the stability restrictions on the reduced root are necessary only near the endpoints provided the function f_y is strictly positive. In other words, the boundary layer behavior is localized to that endpoint and does not affect the behavior of the solution in the interior of the interval. We first state and prove a genral result and then examine several of its consequences in the rest of the section.

THEOREM 5.1. *Assume*

1) *the reduced equation* $f(t,u,u',0)=0$ *has a solution* $u=u(t)$ *of class* $C^{(2)}[-1,1]$ *with* $u(-1)\leqq A$ *and* $u(1)\geqq B$;

2) *the function f is continuous in* (t,y,y',ε) *and of class* $C^{(2)}$ *with respect to* y, y' *in* $\mathscr{E}(-1,1)$: $-1\leqq t\leqq 1, |y-u(t)|\leqq d(t), -1\leqq t\leqq 1, d(t)\geqq A-u(-1), -1\leqq$

$t<-1+\delta$, $d(t)=O(\varepsilon)$, $-1+\delta \leqq t \leqq 1-\delta$, $d(t) \geqq u(1)-B$, $1-\delta<t \leqq 1$, *for* $\delta>0$ *a small constant,* $|y'|<\infty$, $0 \leqq \varepsilon \leqq \varepsilon_0$, $\varepsilon_0>0$; *in* $\mathscr{E}(-1, 1)$, $f_{y'y'}=O(1)$ *as* $|y'| \to \infty$;

3) *there is a constant* $k>0$ *such that* $f_{y'}(t, u(t), u'(t), 0) \leqq -k<0$, $-1 \leqq t \leq -1+\delta$, if $u(-1)<A$, *and* $f_{y'}(t, u(t), u'(t), 0) \geqq k>0$, $1-\delta \leqq t \leqq 1$, *if* $u(1)>B$;

4) *there is a constant* $p>0$ *such that* $f_{y'y'} \geqq p>0$ *in* $\mathscr{E}(-1, -1+\delta)$, *if* $u(-1)<A$; *and* $f_{y'y'} \leqq -p<0$ *in* $\mathscr{E}(1-\delta, 1)$ *if* $u(1)>B$;

5) *there is a constant* $l>0$ *such that* $f_y(t, y, u'(t), 0) \geqq l>0$, $|y-u(t)| \leqq d(t)$, $-1 \leqq t \leqq 1$;

6) $f(t, u(t), u'(t), \varepsilon)=O(\varepsilon)$, $-1 \leqq t \leqq 1$.

Then for each $\varepsilon>0$ *sufficiently small,* $\varepsilon \leqq \varepsilon_0$, *the boundary value problem* (4.1), (4.2) *has a solution* $y=y(t, \varepsilon)$ *with*

$$u(t)-(u(1)-B) \exp[-k_1\varepsilon^{-1}(1-t)]-\varepsilon\bar{\gamma} \leqq y(t, \varepsilon) \leqq u(t)+(A-u(-1)) \exp[-k_1\varepsilon^{-1}(1+t)]+\varepsilon\bar{\gamma}, \qquad -1 \leqq t \leqq 1,$$

for $k_1>0$ *such that* $k_1<k$ *and* k_1-k *is small.*

Proof. Define for t in $[-1, 1]$ and ε in $(0, \varepsilon_0]$,

$$\alpha(t, \varepsilon)=u(t)-(u(1)-B) \exp[-k_1\varepsilon^{-1}(1-t)]-\varepsilon\gamma l^{-1},$$
$$\beta(t, \varepsilon)=u(t)+(A-u(-1)) \exp[-k_1\varepsilon^{-1}(1+t)]+\varepsilon\gamma l^{-1}.$$

We verify explicitly that α satisfies the correct differential inequality; the verification for β proceeds analogously and is omitted. Clearly

$$\begin{aligned}\varepsilon\alpha''-f(t, \alpha, \alpha', \varepsilon)=&\ \varepsilon u''-\varepsilon^{-1}k_1^2(u(1)-B) \exp[-k_1\varepsilon^{-1}(1-t)]-f(t, u, u', \varepsilon)\\ &-f_y(t, \eta, u', \varepsilon)\{-(u(1)-B) \exp[-k_1\varepsilon^{-1}(1-t)]-\varepsilon\gamma l^{-1}\}\\ &-f_{y'}(t, u, u', \varepsilon)\{-k_1\varepsilon^{-1}(u(1)-B) \exp[-k_1\varepsilon^{-1}(1-t)]\}\\ &-f_{yy'}(t, \eta, u', \varepsilon)\{-(u(1)-B) \exp[-k_1\varepsilon^{-1}(1-t)]-\varepsilon\gamma l^{-1}\}\\ &\cdot\{-k_1\varepsilon^{-1}(u(1)-B) \cdot \exp[-k_1\varepsilon^{-1}(1-t)]\}\\ &-\tfrac{1}{2}f_{y'y'}(t, \eta, \eta', \varepsilon)\{k_1\varepsilon^{-1}(u(1)-B) \exp[-k_1\varepsilon^{-1}(1-t)]\}^2\end{aligned}$$

$((t, \eta, \eta', \varepsilon)$ in $\mathscr{E}(-1, 1))$

$$\begin{aligned}\geqq&-\varepsilon|u''|-\varepsilon^{-1}k_1^2(u(1)-B) \exp[-k_1\varepsilon^{-1}(1-t)]\\ &-\varepsilon|\varepsilon^{-1}f(t, u, u', \varepsilon)|+l(u(1)-B) \exp[-k_1\varepsilon^{-1}(1-t)]\\ &+\varepsilon\gamma+f_{y'}(t, u, u', \varepsilon)k_1\varepsilon^{-1}(u(1)-B) \exp[-k_1\varepsilon^{-1}(1-t)]\\ &-L\{(u(1)-B) \exp[-k_1\varepsilon^{-1}(1-t)]+\varepsilon\gamma l^{-1}\}\\ &\cdot k_1\varepsilon^{-1}(u(1)-B) \exp[-k_1\varepsilon^{-1}(1-t)]\\ &-\tfrac{1}{2}f_{y'y'}(t, \eta, \eta', \varepsilon)\{k_1\varepsilon^{-1}(u(1)-B) \exp[-k_1\varepsilon^{-1}(1-t]\}^2.\end{aligned}$$

Here $|f_{yy'}(t, \eta, u', \varepsilon)| \leqq L$. Restricting attention first to the subinterval $[-1, 1-\delta]$, we note that the boundary layer term $\exp[-k_1\varepsilon^{-1}(1-t)]$ is transcendentally small and, consequently, the desired inequality

$$\varepsilon\alpha''-f(t, \alpha, \alpha', \varepsilon) \geqq -\varepsilon|u''|-\varepsilon|\varepsilon^{-1}f(t, u, u', \varepsilon)|-\tau(\varepsilon)+\varepsilon\gamma \geqq 0$$

follows if $\gamma \geqq \max\{|u''|, |\varepsilon^{-1}f(t, u, u', \varepsilon)|, \varepsilon^{-1}\tau(\varepsilon)\}$. Here $\tau(\varepsilon)>0$ represents the contribution of the transcendentally small terms. Finally on the subinterval $[1-\delta, 1]$, we invoke our assumptions about $f_{y'}$ and $f_{y'y'}$ to obtain the inequality

$$\begin{aligned}\varepsilon\alpha''-f(t, \alpha, \alpha', \varepsilon) \geqq &-\varepsilon|u''|-\varepsilon^{-1}k_1^2(u(1)-B)\exp[-k_1\varepsilon^{-1}(1-t)]\\ &-\varepsilon|\varepsilon^{-1}f(t, u, u', \varepsilon)|\\ &+l(u(1)-B)\exp[-k_1\varepsilon^{-1}(1-t)]+\varepsilon\gamma\\ &+kk_1\varepsilon^{-1}(u(1)-B)\exp[-k_1\varepsilon^{-1}(1-t)]\\ &-L\{(u(1)-B)\exp[-k_1\varepsilon^{-1}(1-t)]+\varepsilon\gamma l^{-1}\}\\ &\cdot k_1\varepsilon^{-1}(u(1)-B)\exp[-k_1\varepsilon^{-1}(1-t)]\\ &+\tfrac{1}{2}p\{k_1\varepsilon^{-1}(u(1)-B)\exp[-k_1\varepsilon^{-1}(1-t)]\}^2.\end{aligned}$$

Clearly for ε sufficiently small and $k_1<k, \frac{1}{2}pk_1^2\varepsilon^{-2}>Lk_1\varepsilon^{-1}$ and $k\varepsilon^{-1}>k_1\varepsilon^{-1}+L\gamma l^{-1}$. Thus with the above choice of γ we have the desired inequality on $[1-\delta, 1]$ also. The theorem now follows from Theorem 2.1.

Remark 5.1. The theorem is valid under the weaker assumption that $f_y(t, u(t), u'(t), 0)\geqq l>0$, for t in $(-1+\delta, 1-\delta)$. The proof of this result, which may be of use in concrete applications, is only technically more difficult than that of Theorem 5.1 since it involves the construction of slightly more complicated functions α and β.

Remark 5.2. The analogous result is valid under the restrictions in Remark 4.3 with $u_1=u_2\equiv u$ and $\mathscr{D}=\mathscr{E}$.

Remark 5.3. Theorem 5.1 can be viewed as a generalization of Theorem 4 in [2] to the case of functions f which are quadratically nonlinear in y'.

We consider now two applications of this result. The first case of interest in when the function $f_{y'}(t, u(t), u'(t), \varepsilon)$ is of the order $O(t^{2n}), n\geqq 1$, throughout $[-1, 1]$. If $f_{y'}[u(t)]$ is positive in $[-1, 1]\backslash\{0\}$, we require that $u(-1)=A$; while if $f_{y'}[u(t)]$ is negative there, we require that $u(1)=B$.

In either case we can apply Theorem 5.1 and deduce that a solution of (4.1), (4.2) exists and exhibits boundary layer behavior in a neighborhood of the excluded endpoint. Thus although both $f_{y'}[u(t)]$ and $f_{y'y'}$ vanish at $t=0$, there is no interior layer phenomenon of the type discussed above at $t=0$. The crucial assumption seems to be that $f_{y'}[u(t)]$ does *not* change its algebraic sign as t changes sign.

The second case of interest involves a solution $u=u(t)$ of the reduced equation which cannot satisfy either boundary condition. If the hypotheses of the theorem are satisfied, a solution of the full problem actually exhibits boundary layer behavior at both endpoints. Since $f_{y'}[u(t)]$ is assumed to be negative near $t=-1$ and positive near $t=1$, the root u has the correct stability in terms of boundary layer behavior. Away from $t=\pm 1$, however, its "$f_{y'}$-stability" is immaterial since it can derive stability from f_y. This is the reason for assuming that $f_y(t, y, u', \varepsilon)>0$.

In conclusion we give two examples.

Example 5.1. Consider first the problem

$$\varepsilon y''=t(y')^2+t^2y'+y, \qquad -1<t<1,$$
$$y(-1,\varepsilon)=0, \qquad y(1,\varepsilon)=B\geqq 0.$$

Then the function $u(t)\equiv 0$ trivially satisfies the hypotheses of the theorem, since $u(-1)=y(-1,\varepsilon)$, $f_{y'}[0]=t^2$, and $f_y\equiv 1$. We conclude directly that this problem has a unique solution $y=y(t,\varepsilon)$ for which

$$0\leqq y(t,\varepsilon)\leq B\exp[-k\varepsilon^{-1}(1-t)]+\varepsilon\gamma, \qquad -1\leqq t\leqq 1,$$

with $0<k<1$.

Example 5.2. Finally for the problem

$$\varepsilon y''=-t(y')^2+ty'+y, \qquad -1<t<1,$$
$$y(-1,\varepsilon)=A>0, \qquad y(1,\varepsilon)=B<0,$$

we have again that $u(t)\equiv 0$ satisfies all of the hypotheses of Theorem 5.1. Indeed, $u(-1)<A$, $u(1)>B$, $f_{y'y'}[-1]=1$, $f_{y'y'}[1]=-1$, $f_{y'}[0]=t$ and $f_y\equiv 1$. We conclude immediately that this problem has a unique solution $y=y(t,\varepsilon)$ satisfying

$$B\exp[-(1-\delta)\varepsilon^{-1}(1-t)]-\varepsilon\gamma\leqq y(t,\varepsilon)$$
$$\leqq A\exp[-(1-\delta)\varepsilon^{-1}(1+t)]+\varepsilon\gamma, \qquad -1\leqq t\leqq 1,$$

for $0<\delta<1$.

6. Regular-singular crossings. In [9] we studied the class of problems (4.1), (4.2) under the principal assumption that $f_{y'y'}$ was never zero in its domain of definition. An interesting consequence of this is that if a singular solution of the reduced equation $f(t,u,u',0)=0$ intersects any other solution, then at the point of intersection, the derivatives of the functions also agree. For the problems under consideration here, it is however possible to have an "angular" crossing at $t=0$ between a singular root and another one, since $f_{y'y'}$ vanishes at $t=0$. To be precise, we consider in this section the case in which a regular solution of the reduced equation intersects a singular one at $t=0$ with unequal slopes. It is reasonable to expect a solution of the full problem for small $\varepsilon>0$ to treat this crossing as a Haber–Levison one. In addition, since one of the reduced roots is singular, this solution in general exhibits boundary layer behavior associated with such a root. An illustration of the type of result which is useful in describing such phenomena is provided by the next theorem.

THEOREM 6.1. *Assume*

1) *the reduced equation* $f(t,u,u',0)=0$ *has a regular solution* $u=u_L(t)$ *and a singular solution* $u=u_S(t)$, *defined and of class* $C^{(2)}$ *on* $[-1,0]$ *and* $[0,1]$, *respectively, with* $u_L(-1)=A$, $u_L(0)=u_S(0)$, $u_L'(0)\neq u_S'(0)$ *and* $u_S(1)\leqq B$;

2) *the function* f *is continuous in* (t,y,y',ε) *and of class* $C^{(2)}$ *with respect to* y, y' *in* $\mathcal{M}(-1,1)$: $-1\leqq t\leqq 1$, $|y-u_L(t)|\leqq d(t)$, $-1\leqq t\leqq 0$, $|y-u_S(t)|\leqq d(t)$, $0\leqq t\leqq 1$, *where* $d(t)=O(\varepsilon^{1/2})$, $-1\leqq t\leqq 1-\delta$, *and* $d(t)\geqq|B-u_S(1)|$, $1-\delta<t\leqq 1$, *for* $\delta>0$ *a small constant,* $|y'|<\infty$, $0\leqq\varepsilon\leqq\varepsilon_0$, $\varepsilon_0>0$; *in* $\mathcal{M}(-1,1)$, $f_{y'y'}=O(1)$, *as* $|y'|\to\infty$, *and* $f_{y'y'}(0,\eta,\eta',\varepsilon)=0$, $(0,\eta,\eta',\varepsilon)$ *in* $\mathcal{M}$;

3) $f_{y'}(t,u_L(t),u_L'(t),\varepsilon)\geqq 0$, $-1\leqq t\leqq 0$;

4) *there is a constant* $l>0$ *such that*

$$f_y(t, u_L(t), u_L'(t), 0) \geqq l > 0, \qquad -1 \leqq t \leqq 0;$$

$$f_y(t, y, u_S'(t), 0) \geqq l > 0, \qquad |y - u_S(t)| \leqq d(t), \qquad 0 \leqq t \leqq 1;$$

5) *there is a constant* $p>0$ *such that* $f_{y'y'} \geqq p > 0$ *in* $\mathcal{M}(1-\delta, 1)$, *for* δ *as in* 2);

6)

$$f(t, u_L(t), u_L'(t), \varepsilon) = O(\varepsilon), \qquad -1 \leqq t \leqq 0,$$

$$f(t, u_S(t), u_S'(t), \varepsilon) = O(\varepsilon), \qquad 0 \leqq t \leqq 1.$$

Then for each $\varepsilon > 0$ *sufficiently small,* $\varepsilon \leqq \varepsilon_0$, *the problem* (4.1), (4.2) *has a solution* $y = y(t, \varepsilon)$ *satisfying*

$$u_L(t) - \varepsilon^{1/2}\bar{\gamma} \leqq y(t, \varepsilon) \leqq u_L(t) + \varepsilon^{1/2}\bar{\gamma}, \qquad -1 \leqq t \leqq 0,$$

$$u_S(t) - \varepsilon^{1/2}\bar{\gamma} \leqq y(t, \varepsilon) \leqq u_S(t) + (B - u_S(1)) \exp[-(\varepsilon^{-1}l)^{1/2}(1-t)] + \varepsilon^{1/2}\bar{\gamma}, \qquad 0 \leqq t \leqq 1.$$

Proof. We consider the case in which $u_L'(0) < u_S'(0)$ and for simplicity, we assume that $u_L'' \geqq 0$ and $u_S'' \geqq 0$. Define then for ε in $(0, \varepsilon_1]$, $\varepsilon_1 = \min\{1, \varepsilon_0\}$,

$$\alpha(t, \varepsilon) = \begin{cases} u_L(t), & -1 \leqq t \leqq 0, \\ u_S(t), & 0 \leqq t \leqq 1, \end{cases}$$

$$\beta(t, \varepsilon) = \begin{cases} u_L(t) + (\varepsilon\sigma^{-1})^{1/2}(u_S'(0) - u_L'(0)) + (B - u_S(1)) \exp[-(\varepsilon^{-1}l_1)^{1/2}] \\ \quad + \varepsilon^{1/2}\gamma l^{-1}, & -1 \leqq t \leqq 0, \\ u_S(t) + (\varepsilon\sigma^{-1})^{1/2}(u_S'(0) - u_L'(0)) \exp[-(\varepsilon^{-1}l_1)^{1/2}t] \\ \quad + (B - u_S(1)) \exp[-(\varepsilon^{-1}l)^{1/2}(1-t)] + \varepsilon^{1/2}\gamma l^{-1}, & 0 \leqq t \leqq 1, \end{cases}$$

for $0 < \sigma < l_1 < l$. The verification that these functions satisfy the correct inequalities proceeds without difficulty once we have noted that $D_l\alpha(0) = u_L'(0) < D_r\alpha(0) = u_S'(0)$ and $D_l\beta(0) \geqq D_r\beta(0)$. The case of nonconvex reduced solutions is handled in a similar manner.

It is worth remarking that if u_S satisfies the opposite inequality $u_S(1) > B$ and if the function $f_{y'y'}$ is strictly negative in $\mathcal{M}(1-\delta, 1)$, then a solution of the full problem exists and is *concave* in the boundary layer at $t = 1$. Also a similar result is valid if the first-order stability assumption that f_y is strictly positive is replaced by the higher order stability conditions discussed above. We conclude this section with two examples.

Example 6.1. Consider the problem

$$\varepsilon y'' = y - t(y')^2, \qquad -1 < t < 1,$$

$$y(-1, \varepsilon) = -1, \qquad y(1, \varepsilon) = B \leqq 0.$$

The pertinent roots of the reduced equation are clearly $u_L(t) = t$ and $u_S \equiv 0$, for which $u_L(0) = u_S(0)$ and $u_L'(0) > u_S'(0)$. Since $f_{y'}[u_L(t)] = -2t \geqq 0$, $-1 \leqq t \leqq 0$, $f_{y'y'} =$

$-2t$, $u_S(1) \geqq B$, and $f_y \equiv 1$, we can argue as in the proof of Theorem 6.1 to show that this problem has a unique solution $y = y(t, \varepsilon)$. In addition,

$$\min(t, 0) + B \exp[-\varepsilon^{-1/2}(1-t)] - \varepsilon^{1/2}\gamma \leqq y(t, \varepsilon) \leqq \min(t, 0), \qquad -1 \leqq t \leqq 1.$$

Example 6.2. In this example we illustrate why it is often necessary to examine several of the partial derivatives $\partial_y^j f[u_S]$, $j \geqq 1$. Consider

$$\varepsilon y'' = y^3 - t^3(y')^2, \qquad -1 < t < 1,$$
$$y(-1, \varepsilon) = A < 0, \qquad y(1, \varepsilon) = B \leqq 0.$$

It is not difficult to see that $u_L(t) = t((-t)^{1/2}((-A)^{1/2} - 1) + 1)^{-2}$, $-1 \leqq t \leqq 0$, and $u_S \equiv 0$, $0 \leqq t \leqq 1$, form a stable pair which have an angular crossing at $t = 0$. Look now at $f_y = 3y^2$; clearly, $f_y[0] = f_{yy}[0] = 0$, but $f_{yyy} \equiv 6$. Consequently the proof of Theorem 6.1 can be adapted to show that this problem has a solution $y = y(t, \varepsilon)$. In addition,

$$u(t) + B(1 + (2\varepsilon)^{-1/2}(1-t))^{-1} - \varepsilon^{1/6}\gamma \leqq y(t, \varepsilon) \leqq u(t) + \varepsilon^{1/6}\gamma, \qquad -1 \leqq t \leqq 1,$$

for

$$u(t) = \begin{cases} u_L(t), & -1 \leqq t \leqq 0, \\ 0, & 0 \leqq t \leqq 1. \end{cases}$$

7. Related phenomena. We conclude by discussing briefly several other phenomena which solutions of problems with generalized turning points can exhibit. Two of these phenomena are treated in detail in [9] and for this reason, we only indicate here the possible extensions of the theory given above.

The first remark concerns conditions under which there exist solutions of the full problem exhibiting boundary layer behavior at one (or both) of the endpoints. In § 5 we gave one set of sufficient conditions for such behavior. However it is more common for a single solution of the reduced equation to have fixed stability properties throughout $[-1, 1]$. For example, in the case of a reduced root $u = u_L(t)$, one can assume $u_L(-1) = y(-1, \varepsilon)$ and $f_{y'}[u_L(t)] \geqq k > 0$, $-1 \leqq t \leqq 1$. Then if the value of $u_L(1)$ relative to $y(1, \varepsilon)$ is compatible with the sign of $f_{y'y'}$ near $t = 1$ (cf. § 5), it is not difficult to show that a solution $y = y(t, \varepsilon)$ of the full problem exists and satisfies $y(t, \varepsilon) = u_L(t) + O(|y(1, \varepsilon) - u_L(1)| \exp[-k\varepsilon^{-1}(1-t)]) + O(\varepsilon)$, $-1 \leqq t \leqq 1$. Similarly, if the reduced equation has a solution $u = u_R(t)$ with $u_R(1) = y(1, \varepsilon)$ and $f_{y'}[u_R(t)] \leqq -k < 0$, $-1 \leqq t \leqq 1$, then under the appropriate sign restrictions at $t = -1$, a solution $y(t, \varepsilon)$ exists and satisfies $y(t, \varepsilon) = u_R(t) + O(|y(-1, \varepsilon) - u_R(-1)| \exp[-k\varepsilon^{-1}(1+t)]) + O(\varepsilon)$, $-1 \leqq t \leqq 1$. The global stability of these roots u_L, u_R makes the above results valid without any positivity restriction on $f_y(t, y, u'_{L,R}, 0)$. In the same vein, it is possible to generate boundary layer behavior, in general at both endpoints, by considering globally stable singular reduced roots. Many of these results which are treated extensively in [9, § 3] for $f_{y'y'}$ never zero are valid in the case of functions with generalized turning points in $(-1, 1)$.

This brings us to the question of what happens when the generalized turning point coincides with an endpoint. Under the appropriate assumptions, a solution of the full problem exhibits boundary layer behavior at such a point (say $t = -1$)

which is described by a term of the form $\exp[\varepsilon^{-1}\int_{-1}^{t} h(s)\,ds]$ where $f_{y'} \leqq h(t) \leqq 0$ near $t=-1$. If $f_{y'y'}$ vanishes at $t=-1$ but is not identically zero near $t=-1$, any solution of the full problem for small $\varepsilon>0$ is influenced decisively by the nature of $f_{y'y'}$. It is then very often possible to combine the theory described above in § 3 with that in [9, § 3] to study the behavior of solutions in the neighborhood of an endpoint at which $f_{y'y'}$ is zero. In this regard, see also [7, § 6].

Finally we note that the interior crossing theory discussed above consists of only two examples of such behavior—Haber–Levinson crossings and angular singular-singular crossings. Another important type of internal crossing, the smooth regular-singular crossing, occurs provided $f_{y'y'}$ is not zero at the crossing point and is discussed extensively in [9, § 6]. In particular, this theory is applicable to problems with generalized turning points if the point of the regular-singular crossing is different from the turning point. These three basic types of internal crossings can occur in various combinations in the solution of an actual problem. That is, the reduced equation may have several regular and singular solutions which cross each other at a number of points inside the interval. To determine the asymptotic behavior of a solution of the full problem, one then examines the nature of the reduced roots at each crossing point and applies the theory given above and in [9]. This qualitative picture of the solution for small ε can be given a quantitative formulation by constructing bounding functions α and β which incorporate such behavior into their boundary and interior layer terms.

Acknowledgment. The author is grateful to the Mathematics Department of the University of Wisconsin—Madison for partially supporting this work. He also wishes to express his warmest thanks to Professor Wolfgang R. Wasow for inviting him to Wisconsin.

REFERENCES

[1] YU. P. BOGLAEV, *The two-point problem for a class of ordinary differential equations with a small parameter coefficient of the derivative*, USSR Computational Math. and Math. Phys., 10 (1970), no. 4, pp. 191–204.

[2] N. I. BRIŠ, *On boundary value problems for the equation $\varepsilon y''=f(x, y, y')$ for small ε*, Dokl. Akad. Nauk SSSR, 95 (1954), pp. 429–432.

[3] F. W. DORR, *Some examples of singular perturbation problems with turning points*, this Journal, 1 (1970), pp. 141–146.

[4] F. W. DORR, S. V. PARTER AND L. F. SHAMPINE, *Applications of the maximum principle to singular perturbation problems*, SIAM Rev., 15 (1973), pp. 43–88.

[5] S. HABER AND N. LEVINSON, *A boundary value problem for a singularly perturbed differential equation*, Proc. Amer. Math. Soc., 6 (1955), pp. 866–872.

[6] P. HABETS AND M. LALOY, *Etude de problemes aux limites par la methode des sur-et sous-solutions*, Lecture Notes, Catholic University of Louvain, Louvain, Belgium, 1974.

[7] F. A. HOWES, *Singularly perturbed nonlinear boundary value problems with turning points*, this Journal, 6 (1975), pp. 644–660.

[8] ———, *A class of boundary value problems whose solutions possess angular limiting behavior*, Rocky Mountain J. Math., 6 (1976), pp. 591–607.

[9] ———, *The asymptotic solution of a class of singularly perturbed nonlinear boundary value problem via differential inequalities*, this Journal, 9 (1978), pp. 215–249.

[10] ———, *The asymptotic solution of boundary value problems with angular limiting solutions*, Trans. Amer. Math. Soc., to appear.

[11] L. K. JACKSON, *Subfunctions and second order ordinary differential inequalities*, Advances in Math., 2 (1968), pp. 307–363.

[12] M. NAGUMO, *Über die Differentialgleichung* $y''=f(x, y, y')$, Proc. Phys. Math. Soc. Japan, 19 (1937), pp. 861–866.

[13] R. E. O'MALLEY, JR., *On boundary value problems for a singularly perturbed differential equation with a turning point*, this Journal, 1 (1970), pp. 479–490.

[14] ———, *On singular perturbation problems with interior nonuniformities*, J. Math. Mech., 19 (1970), pp. 1103–1112.

[15] ———, *Introduction to Singular Perturbations*, Academic Press, New York, 1974.

[16] A. B. VASIL'EVA, *Asymptotic behavior of solutions to certain problems involving nonlinear differential equations containing a small parameter multiplying the highest derivatives*, Russian Math. Surveys, 18 (1963), pp. 13–84.

[17] M. I. VISHIK AND L. A. LIUSTERNIK, *Initial jump for nonlinear differential equations containing a small parameter*, Sov. Math. Dokl., 1 (1960), pp. 749–752.

[18] W. R. WASOW, *Asymptotic Expansions for Ordinary Differential Equations*, Interscience, New York, 1965.

SIAM J. MATH. ANAL.
Vol. 9, No. 2, April 1978
Copyright © 1978 Society for Industrial and Applied Mathematics

RATIONAL APPROXIMANT BOUNDS FOR A CLASS OF TWO-VARIABLE STIELTJES FUNCTIONS*

M. F. BARNSLEY† AND P. D. ROBINSON†

Abstract. Dual variational principles are used to impose upper and lower bounds on functions $F(w, z)$ which can be written $F(w, z)=\int_0^\infty \int_0^\infty d\sigma(s, t)/(1+ws+zt)$, where $\sigma(s, t)$ is bounded, monotone nondecreasing in s for $0\leqq s<\infty$ and $0\leqq t<\infty$. Following a procedure analogous to one which has been used for single-variable Stieltjes functions, trial vectors are chosen so as to yield bounds on $F(w, z)$ for all $0\leqq w<\infty$, $0\leqq z<\infty$, in the form of rational approximants involving only an initial set of coefficients in the formal double series expansion $F(w, z)\sim\sum_{m=0}^\infty \sum_{n=0}^\infty (-1)^{m+n}F_{m,n}w^m z^n$. It is proved that the approximants display certain matching properties with the formal expansion of $F(w, z)$ when their series expansions are taken. Furthermore, it is established that the bounds obtained improve upon those found by treating $F(w, z)$ as a one-variable Stieltjes function and forming Padé approximants using a comparable set of given coefficients. Among the approximants defined here are those suggested by Alabiso and Butera on moment-theoretic grounds. Some applications of the theory, together with simple examples, are given.

1. Introduction. It has been shown that the theory of Padé approximants (PA's) for Stieltjes functions can be evolved in a natural way by taking dual variational principles as a starting point [1], [2], [3]. Here we consider two-variable Stieltjes functions expressible in the form

$$F(w, z)=\int_0^\infty \int_0^\infty \frac{d\sigma(s, t)}{(1+ws+zt)}; \tag{1.1}$$

where $\sigma(s, t)$ on $0\leqq s<\infty$, $0\leqq t<\infty$, is bounded, monotone nondecreasing in s for fixed t, and monotone nondecreasing in t for fixed s; and where the variables w and z both belong to the complex plane cut from $-\infty$ to $0-$. This function has the formal double-series expansion

$$F(w, z)\sim \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} (-1)^{m+n}\binom{m+n}{n} F_{m,n}w^m z^n \tag{1.2}$$

where

$$F_{m,n}=\int_0^\infty \int_0^\infty s^m t^n \, d\sigma(s, t), \qquad m, n=0, 1, \cdots. \tag{1.3}$$

We suppose that we know an initial set of coefficients occurring in (1.2), say

$$F_{0,0}, F_{1,0}, F_{0,1}, F_{2,0}, F_{1,1}, F_{0,2}, F_{2,1}, \cdots, F_{j,k}, \tag{1.4}$$

and assume that these are finite. Then we ask what bounds can be imposed on $F(w, z)$ when $0\leqq w<\infty$, $0\leqq z<\infty$, on the basis of the given information.

In § 2 we give a dual pair of variational functionals which impose upper and lower bounds on $F(w, z)$ for all real positive w and z. We then observe the general "rational approximant" structure of the optimal bounds, obtained when linear variational trial vectors are used, and note what sets of information are required. In § 3 we choose variational basis sets as large as possible, but in such a way as to ensure that the only information needed for the construction of the bounds is an initial set of coefficients

* Received by the editors April 23, 1976.

† School of Mathematics, University of Bradford, Bradford, England BD7 1DP. This work was supported in part by the Science Research Council.

occurring in (1.2): we denote these bounds by

$$J(N, R, S) \leqq F(w, z) \leqq G(N, R, S), \qquad w \geqq 0, \quad z \geqq 0, \tag{1.5}$$

where (N, R, S) indexes the $F_{m,n}$'s which are used. Among the approximants obtained here are those defined by Alabiso and Butera [4] on moment theoretic grounds, appropriate for two-variable *extended* Stieltjes functions. In § 4 we use the variationally optimal nature of $J(N, R, S)$ and $G(N, R, S)$ to establish the matching properties

$$F(w, z) - J(N, R, S) \sim \text{terms of order } w^p z^q, \qquad p+q \geqq 2N+2, \tag{1.6}$$

and

$$F(w, z) - G(N, R, S) \sim \text{terms of order } w^p z^q, \qquad p+q \geqq 2N+3. \tag{1.7}$$

The derivation of the above results generalizes previous work [3] on the characterization of the $[N/(N+1)]$, $[N+1/(N+1)]$ optimal pair of PA's for one-variable Stieltjes functions.

An alternative way of imposing upper and lower bounds on $F(w, z)$ by utilizing only an initial set of coefficients in the formal series (1.2) is by forming what are essentially PA's for a single-variable Stieltjes function. That is, one replaces w and z in (1.1) by λw and λz respectively so that, for fixed positive w and z, (1.2) becomes a series of Stieltjes in λ. Applying the usual PA theory [5], we obtain the bounds

$$[N/(N+1)](w, z) \leqq F(w, z) \leqq [N+1/(N+1)](w, z), \qquad w \geqq 0, \quad z \geqq 0, \tag{1.8}$$

by putting $\lambda = 1$ in the pair of PA's $[N/(N+1)]$ and $[N+1/(N+1)]$ for $F(\lambda w, \lambda z)$ as a function of λ. In § 5 we show that the bounds (1.5) are in general superior to (1.8), whilst the sets of information required are often identical.

In § 6 we consider applications of the theory to the evaluation of bounds on special mathematical functions of the form (1.1), to the improvement of the usual PA bounds for one-variable Stieltjes functions when half-integer moments are available, and to the summation of double perturbation series. Simple numerical examples are given.

We stress here that it is not the purpose of this paper to present an alternative set of two-variable rational approximants to the Chisholm approximants [6] and their relatives [7]. Approximants of the latter types have been designed so that they generalize various key properties of one-variable PA's such as homographic, reciprocal, and unitary invariances [8], [9]. As such, it is expected [10] that they will cope in a general way with problems of analytic continuation of functions specified only by their double-series expansions. We are here concerned for the most part with functions of the special form (1.1), and our emphasis is on bounding properties rather than convergence. Nonetheless, we do go some way towards answering the question, posed by Chisholm [6], of how two-variable rational approximants should be defined in order that their relationship with the two-dimensional moment problem be analogous to the relationship of PA's to Stieltjes functions.

2. Dual variational bounds for $F(w, z)$, $w \geqq 0$, $z \geqq 0$. We begin, following Alababiso and Butera [4], by writing $F(w, z)$ in the form

$$F(w, z) = \langle f, (1 + wA + zB)^{-1} f \rangle, \qquad w \geqq 0, \quad z \geqq 0, \tag{2.1}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in a real Hilbert space $\mathscr{h}$, $f \in \mathscr{h}$, and A and B are a commuting pair of positive self-adjoint linear operators in $\mathscr{h}$ with domains $\mathscr{D}(A)$ and $\mathscr{D}(B)$ respectively. Specifically, $\mathscr{h}$ is the Hilbert space of real functions $\mathscr{h}(s, t)$, $0 \leqq s < \infty$, $0 \leqq t < \infty$, which are square integrable with respect to the measure $\sigma(s, t)$, and the inner product between $\hbar_1 \in \hbar$ and $\hbar_2 \in \hbar$ is

$$\langle \hbar_1, \hbar_2 \rangle = \int_0^\infty \int_0^\infty \hbar_1(s,t)\hbar_2(s,t)\, d\sigma(s,t). \tag{2.2}$$

A and B are the linear operators which multiply by s and t respectively, so that for example

$$A\eta = A\eta(s,t) = s\eta(s,t) = \tilde{\eta}, \qquad \text{for all } \eta \in \mathscr{D}(A), \tag{2.3}$$

where $\mathscr{D}(A)$ is precisely the set of $\eta \in \hbar$ such that $\tilde{\eta} \in \hbar$. Then, since both A and B are positive, the equation

$$(1 + wA + zB)\phi = f, \qquad w \geqq 0, \quad z \geqq 0, \tag{2.4}$$

possesses a unique solution ϕ for each $f \in \hbar$. With

$$f = f(s,t) = 1 \quad \text{for all } 0 \leqq s < \infty, \quad 0 \leqq t < \infty, \tag{2.5}$$

which belongs to $\hbar$ by virtue of the boundedness of $\sigma(s,t)$, we have in particular,

$$\langle \phi, f \rangle = \langle f, (1 + wA + zB)^{-1} \rangle = \int_0^\infty \int_0^\infty \frac{d\sigma(s,t)}{(1 + ws + zt)} = F(w,z). \tag{2.6}$$

We observe that the coefficients $F_{m,n}$ occurring in (1.2) can be expressed as

$$F_{m,n} = \langle f, A^m B^n f \rangle, \qquad m, n = 0, 1, 2, \cdots. \tag{2.7}$$

When both $w \geqq 0$ and $z \geqq 0$, equation (2.4) takes the form

$$(1 + L)\phi = f, \qquad f \in \hbar, \tag{2.8}$$

where the self-adjoint linear operator

$$L = wA + zB, \qquad \mathscr{D}(L) = \mathscr{D}(A) \cap \mathscr{D}(B), \tag{2.9}$$

satisfies

$$\langle \Theta, L\Theta \rangle \geqq 0 \quad \text{for all } \Theta \in \mathscr{D}(L). \tag{2.10}$$

On applying the theory of dual variational principles [11] to (2.8) one obtains the seemingly most elementary pair of complementary functionals [3]

$$J(\Phi) = -\langle \Phi, (1 + L)\Phi \rangle + 2\langle \Phi, f \rangle, \qquad \Phi \in \mathscr{D}(L), \tag{2.11}$$

and

$$G(\Psi) = \langle f, f \rangle + \langle \Psi, L(1 + L)\Psi \rangle - 2\langle L\Psi, f \rangle, \qquad \Psi \in \mathscr{D}(L), \tag{2.12}$$

where Φ and Ψ are trial vectors. Each functional provides a variational approximation to the quantity $\langle \phi, f \rangle$, for we have

$$J(\Phi) - \langle \phi, f \rangle = -\langle \delta\Phi, (1 + L)\delta\Phi \rangle, \tag{2.13}$$

and

$$G(\Psi) - \langle \phi, f \rangle = \langle \delta\Psi, L(1 + L)\delta\Psi \rangle, \tag{2.14}$$

where $\delta\Phi = \Phi - \phi$, and $\delta\Psi = \Psi - \phi$, Using (2.10) in (2.13) and (2.14), we obtain the complementary bounding properties

$$J(\Phi) \leqq \langle \phi, f \rangle \leqq G(\Psi) \quad \text{for each } w \geqq 0, \quad z \geqq 0. \tag{2.15}$$

Thus, in view of the Hilbert space representation (2.6) for $F(w, z)$, we have at our disposal a mechanism for imposing upper and lower bounds on two-variable Stieltjes functions when $w \geqq 0$, $z \geqq 0$.

We examine here the general structure of bounds on $F(w, z) = \langle \phi, f \rangle$, derived from J and G by means of optimization using linear variational trial vectors

$$\Phi, \Psi = \sum_{n=1}^{N} \alpha_n \Theta_n, \quad \text{where } \alpha_n \in \mathbb{R}, \qquad \Theta_n \in \mathscr{D}(L), \quad n = 1, 2, \cdots, N. \tag{2.16}$$

We will use P to denote the projection operator on $\mathscr{h}$ corresponding to the subspace spanned by $\{\Theta_n : n = 1, 2, \cdots, N\}$. Consider J first. On requiring $J(\Phi)$ to be stationary with respect to variations in the α_n's, and hence maximal, it is found that the optimal choice for Φ is the unique solution Φ_{opt} of the problem

$$\begin{aligned} &P(1+L)P\Phi_{\text{opt}} = Pf, \\ &P\Phi_{\text{opt}} = \Phi_{\text{opt}}, \end{aligned} \tag{2.17}$$

and correspondingly

$$J(\Phi_{\text{opt}}) = \langle \Phi_{\text{opt}}, f \rangle$$

$$= -\frac{\begin{vmatrix} 0 & \Theta_1, f\rangle & \cdots & \langle \Theta_N, f \rangle \\ \langle \Theta_1, f \rangle & \langle \Theta_1, (1+L)\Theta_1 \rangle & \cdots & \langle \Theta_1, (1+L)\Theta_N \rangle \\ \langle \Theta_2, f \rangle & \langle \Theta_2, (1+L)\Theta_1 \rangle & \cdots & \langle \Theta_2, (1+L)\Theta_N \rangle \\ \vdots & & & \\ \langle \Theta_N, f \rangle & \langle \Theta_N, (1+L)\Theta_1 \rangle & \cdots & \langle \Theta_N, (1+L)\Theta_N \rangle \end{vmatrix}}{\begin{vmatrix} \langle \Theta_1, (1+L)\Theta_1 \rangle & \cdots & \langle \Theta_1, (1+L)\Theta_N \rangle \\ \langle \Theta_2, (1+L)\Theta_1 \rangle & \cdots & \langle \Theta_2, (1+L)\Theta_N \rangle \\ \vdots & & \vdots \\ \langle \Theta_N, (1+L)\Theta_1 \rangle & \cdots & \langle \Theta_N, (1+L)\Theta_N \rangle \end{vmatrix}}. \tag{2.18}$$

Here we have assumed that the set $\{\Theta_1, \Theta_2, \cdots, \Theta_N\}$ is linearly independent—if it isn't then the largest subset of the Θ_n's which *is* linearly independent replaces $\{\Theta_1, \Theta_2, \cdots, \Theta_N\}$ in (2.18). In practice this means that if any row in either numerator or denominator of (2.18) is linearly dependent on its predecessors then it must be omitted along with the corresponding column in *both* numerator *and* denominator. The usefulness of this rule will become clear later on.

In order to evaluate $J(\Phi_{\text{opt}})$ we need to know the following sets of information:

(i) the values of the inner products $\langle \Theta_n, f \rangle$, $n = 1, 2, \cdots, N$;

(ii) the values of the matrix elements $\langle \Theta_m, (1+L)\Theta_n \rangle$, $m, n = 1, 2, \cdots, N$.

These can be readily interpreted as various integrals with respect to the measure $\sigma(s, t)$.

In a similar manner, choosing Ψ as in (2.16) and requiring $G(\Psi)$ to be stationary with respect to variations in the α_n's, we find that the optimal choice for Ψ is the unique solution Ψ_{opt} of the problem

$$\begin{aligned} &PL(1+L)P\Psi_{\text{opt}} = PLf, \\ &P\Psi_{\text{opt}} = \Psi_{\text{opt}}, \end{aligned} \tag{2.19}$$

and correspondingly

$$G(\Phi_{\text{opt}})=\langle f-L\Psi_{\text{opt}},f\rangle$$

$$(2.20)\qquad =\langle f,f\rangle+\frac{\begin{vmatrix} 0 & \langle\Theta_1,Lf\rangle & \cdots & \langle\Theta_N,Lf\rangle \\ \langle\Theta_1,Lf\rangle & \langle\Theta_1,L(1+L)\Theta_1\rangle & \cdots & \langle\Theta_1,L(1+L)\Theta_N\rangle \\ \langle\Theta_2,Lf\rangle & \langle\Theta_2,L(1+L)\Theta_1\rangle & \cdots & \langle\Theta_2,L(1+L)\Theta_N\rangle \\ \vdots & \vdots & & \vdots \\ \langle\Theta_N,Lf\rangle & \langle\Theta_N,L(1+L)\Theta_1\rangle & \cdots & \langle\Theta_N,L(1+L)\Theta_N\rangle \end{vmatrix}}{\begin{vmatrix} \langle\Theta_1,L(1+L)\Theta_1\rangle & \cdots & \langle\Theta_1,L(1+L)\Theta_N\rangle \\ \langle\Theta_2,L(1+L)\Theta_1\rangle & \cdots & \langle\Theta_2,L(1+L)\Theta_N\rangle \\ \vdots & & \vdots \\ \langle\Theta_N,L(1+L)\Theta_1\rangle & \cdots & \langle\Theta_N,L(1+L)\Theta_N\rangle \end{vmatrix}}$$

Here we assume that the row vectors occurring in the dividing determinant are linearly independent—if they are not then in practice we follow the same rule as we do in the case of (2.18). In order to evaluate $G(\Psi_{\text{opt}})$ we need to know the following sets of information:

(iii) the value of $\langle f,f\rangle$;

(iv) the values of the inner products $\langle\Theta_n,Lf\rangle$, $n=1,2,\cdots,N$;

(v) the values of the matrix elements $\langle\Theta_m,L(1+L)\Theta_n\rangle$, $m,n=1,2,\cdots,N$.

Provided that the basis functions $\{\Theta_1,\Theta_2,\cdots,\Theta_N\}$ are themselves independent of w and z, we see from (2.18) that we can always write

$$(2.21)\qquad J(\Phi_{\text{opt}})=\sum_{\substack{n+m=0\\ n\geqq 0,m\geqq 0}}^{n+m=N-1} a_{m,n}w^m z^n \Big/ \sum_{\substack{n+m=0\\ n\geqq 0,m\geqq 0}}^{n+m=N} b_{m,n}w^m z^n$$

where the $a_{m,n}$'s and $b_{m,n}$'s are real constant coefficients. When $w=\lambda z$ with λ a constant, $J(\Phi_{\text{opt}})$ reduces to a polynomial of degree $(N-1)$ in z divided by one of degree N. Similarly, from (2.20), we find that we can write

$$(2.22)\qquad G(\Psi_{\text{opt}})=\sum_{\substack{n+m=N\\ n\geqq 0,m\geqq 0}}^{n+m=2N} c_{m,n}w^m z^n \Big/ \sum_{\substack{n+m=N\\ n\geqq 0,m\geqq 0}}^{n+m=2N} d_{m,n}w^m z^n$$

where the $c_{m,n}$'s and $d_{m,n}$'s are real constant coefficients. This involves a "doubling-up" of degrees in the numerator and denominator, but we observe that when $w=\lambda z$ with λ a constant $G(\Psi_{\text{opt}})$ reduces to a polynomial of degree N in z divided by another one of degree N, which makes it appear more consistent with $J(\Phi_{\text{opt}})$.

By putting $w=\lambda z$ with $\lambda\geqq 0$ we can obtain a further understanding of the structures of $J(\Phi_{\text{opt}})$ and $G(\Psi_{\text{opt}})$. In this case

$$(2.23)\qquad L=z(\lambda A+B)$$

and we can without loss of generality replace the Θ_n's by $\tilde{\Theta}_n$'s, spanning the same subspace, such that the $\tilde{\Theta}_n$'s are orthogonal, normalized, and chosen to diagonalize the positive self-adjoint operator L. That is

$$(2.24)\qquad \langle\tilde{\Theta}_i,\tilde{\Theta}_j\rangle=\delta_{ij},\qquad \langle\tilde{\Theta}_i,L\tilde{\Theta}_j\rangle=z\varepsilon_i(\lambda)\delta_{ij}\qquad i,j=1,2,\cdots,N,$$

where δ_{ij} is Kronecker's delta and the $\varepsilon_i(\lambda)$'s are such that

$$0 \leqq \varepsilon_1(\lambda) \leqq \varepsilon_2(\lambda) \leqq \cdots \leqq \varepsilon_N(\lambda). \tag{2.25}$$

We note that if the eigenvalues of $\lambda A + B$ are discrete, which in general corresponds to $\sigma(s, t)$ being piecewise constant, then the $\varepsilon_i(\lambda)$'s are upper bounds to the first N eigenvalues of $\lambda A + B$ taken in order. One way of seeing this is to consider the Rayleigh–Ritz variational principle for the eigenvalues of $\lambda A + B$ with $\{\tilde{\Theta}_1, \tilde{\Theta}_2, \cdots, \tilde{\Theta}_N\}$ as a linear variational basis set (see [12] for example). Substituting into (2.18) and carrying out a partial fractions expansion we find

$$J(\Phi_{\text{opt}}) = \sum_{n=1}^{N} \frac{|\langle \tilde{\Theta}_n, f\rangle|^2}{(1 + z\varepsilon_n(\lambda))} \quad \text{when } w = \lambda z, \text{ and } \lambda \geqq 0. \tag{2.26}$$

Here the nonnegative real numbers $|\langle \tilde{\Theta}_n, f\rangle|^2$ depend on λ but are independent of z. If there are linear dependences among the Θ_n's then the number of terms in the summation is correspondingly decreased. Applying similar reasoning to $G(\Psi_{\text{opt}})$ we find that in general it can be expressed as

$$G(\Psi_{\text{opt}}) = \langle f, f\rangle - \sum_{n=1}^{N} \frac{z\nu_n(\lambda)}{(1 + z\tilde{\varepsilon}_n(\lambda))} \quad \text{when } w = \lambda z, \text{ and } \lambda \geqq 0, \tag{2.27}$$

where the $\nu_n(\lambda)$'s are nonnegative real numbers and the $\tilde{\varepsilon}_n(\lambda)$'s have the same relationship (2.25) as the $\varepsilon_n(\lambda)$'s.

In summary, we have the complementary bounds

$$J(\Phi_{\text{opt}}) \leqq F(w, z) \leqq G(\Psi_{\text{opt}}), \qquad w \geqq 0, \quad z \geqq 0, \tag{2.28}$$

where $J(\Phi_{\text{opt}})$ and $G(\Psi_{\text{opt}})$ are rational functions, given explicitly by (2.18) and (2.20). In the next section we show how the basis functions $\{\Theta_1, \Theta_2, \cdots, \Theta_N\}$ can be chosen such that the requisite information, namely (i), (ii), (iii), (iv) and (v), consists simply of sets of coefficients $F_{m,n}$ taken from the beginning of the double series expansion (1.2).

3. Optimal bounds for $F(w, z)$ which use only sets of $F_{m,n}$'s. When $w = 0$, $F(w, z)$ reduces to a Stieltjes function in the single variable z. In this case it has been shown [3] that the optimal pair, $[N/(N+1)]$, and $[N+1/(N+1)]$, of PA's for $F(0, z)$ are given by $J(\Phi_{\text{opt}})$, and $G(\Psi_{\text{opt}})$, respectively, when the basis vectors are chosen to be $\Theta_1 = f$, $\Theta_n = B^{n-1}f$, $n = 2, 3, \cdots, N$. That is, the basis set consists of the first N vectors occurring in the formal expansion

$$\phi_{w=0} \sim f + \sum_{n=1}^{\infty} (-z)^n B^n f. \tag{3.1}$$

Proceeding analogously in the two-variable case, it seems natural to choose as basis set the "first" N vectors in the formal expansion

$$\phi \sim f + \sum_{\substack{m+n=1 \\ m\geqq 0,\, n\geqq 0}}^{\infty} (-w)^m(-z)^n \binom{m+n}{n} A^m B^n f. \tag{3.2}$$

Accordingly, we take as basis set all of the vectors in (α) and (β), below:

(α) f, together with all vectors of the form $A^m B^n f$ where $m + n = 1, 2, \cdots, N$, with m, n, and N being nonnegative integers;

(β) $\{A^R B^{N+1-R} f, A^{R-1} B^{N+2-R}, \cdots, A^{N+1-S} B^S B^S f\}$ where R and S are integers such that $R + S \geqq N$, $0 \leqq R \leqq N+1$, and $0 \leqq S \leqq N+1$. When $R + S = N$ we understand that there are no vectors in (β).

This basis set can be indexed by the triple (N, R, S), and the total number of vectors is

$$\mathcal{N} = \mathcal{N}(N, R, S) = \tfrac{1}{2}(N+1)(N+2) + (R+S-N). \tag{3.3}$$

Assuming that they have been ordered, we denote them by

$$\{\Theta_1, \Theta_2, \cdots, \Theta_{\mathcal{N}}\} \tag{3.4}$$

exactly as in § 2. With reference to this basis set we will write

$$J(\Phi_{\text{opt}}) = J(N, R, S), \qquad G(\Psi_{\text{opt}}) = G(N, R, S). \tag{3.5}$$

Using the relation (2.7) and the fact that A and B commute, we find from (i) and (ii) that in order to construct $J(N, R, S)$ we need to know the coefficients (see Fig. 1)

$$\{F_{m,n} : m+n = 0, 1, \cdots, 2N+1,\ m \geqq 0,\ n \geqq 0\}, \tag{3.6}$$

$$\{F_{N+1+R,N+1-R}, F_{N+R,N+2-R}, \cdots, F_{N+1-S,N+1+S}\}, \tag{3.7}$$

and

$$\{F_{2R+1,2N+2-2R}, F_{2R,2N+3-R}, \cdots, F_{2N+2-2S,2S+1}\}. \tag{3.8}$$

The latter two sets must be taken to be empty when $R+S=N$.

Similarly, from (iii), (iv), and (v), we find that in order to construct $G(N, R, S)$ we need to know the coefficients (see Fig. 2)

$$\{F_{m,n} : m+n = 0, 1, \cdots, 2N+2,\ m \geqq 0,\ n \geqq 0\}, \tag{3.9}$$

$$\{F_{N+2+R,N+1-R}, F_{N+1-R,N+2-R}, \cdots, F_{N+1-S,N+2+S}\} \tag{3.10}$$

and

$$\{F_{2R+2,2N+2-2R}, F_{2R+1,2N+3-2R}, \cdots, F_{2N+2-2S,2S+2}\}, \tag{3.11}$$

where the latter two sets must be taken to be empty when $R+S=N$.

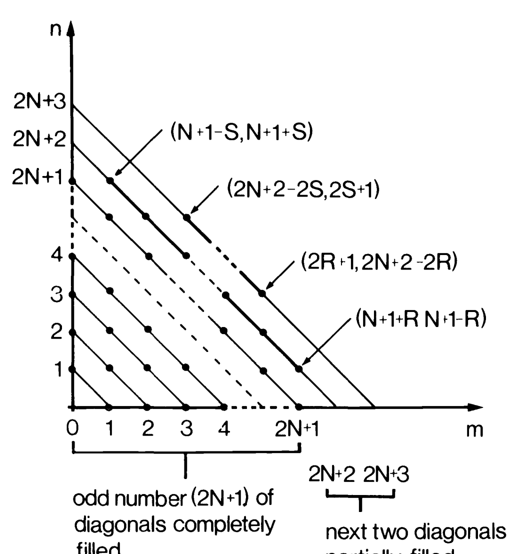

FIG. 1. *Index pairs (m, n) corresponding to the $F_{m,n}$'s required for the construction of $J(N, R, S)$.*

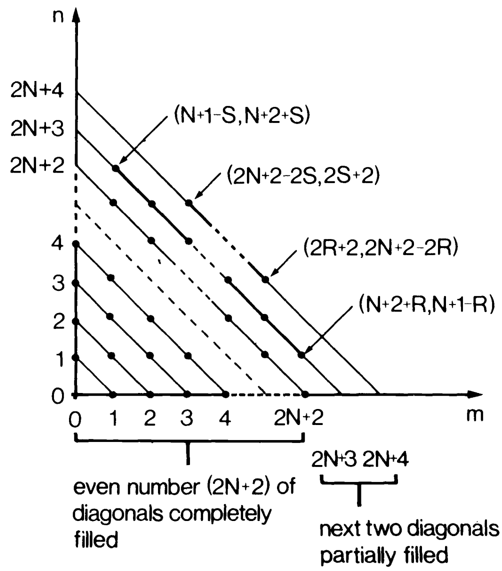

FIG. 2. *Index pairs (m, n) corresponding to the $F_{m,n}$'s required for the construction of $G(N, R, S)$.*

Given that one knows the coefficients (3.6), (3.7), and (3.8), for some allowed (N, R, S), the corresponding $J(N, R, S)$ can be obtained by substituting for the basis functions $\{\Theta_1, \Theta_2, \cdots, \Theta_{\mathcal{N}}\}$ described in (α) and (β), in (2.18). We obtain for example

$$J(1,1,1) = -\begin{vmatrix} 0 & \langle f,f\rangle & \langle Af,f\rangle & \langle Bf,f\rangle & \langle ABf,f\rangle \\ \langle f,f\rangle & \langle f,(1+wA+zB)f\rangle & \langle f,(1+wA+zB)Af\rangle & \langle f,(1+wA+zB)Bf\rangle & \langle f,(1+wA+zB)ABf\rangle \\ \langle Af,f\rangle & \langle Af,(1+wA+zB)f\rangle & \langle Af,(1+wA+zB)Af\rangle & \langle Af,(1+wA+zB)Bf\rangle & \langle Af,(1+wA+zB)ABf\rangle \\ \langle Bf,f\rangle & \langle Bf,(1+wA+zB)f\rangle & \langle Bf,(1+wA+zB)Af\rangle & \langle Bf,(1+wA+zB)Bf\rangle & \langle Bf,(1+wA+zB)ABf\rangle \\ \langle ABf,f\rangle & \langle ABf,(1+wA+zB)f\rangle & \langle ABf,(1+wA+zB)Af\rangle & \langle ABf,(1+wA+zB)Bf\rangle & \langle ABf,(1+wA+zB)ABf\rangle \end{vmatrix}$$

(3.12) $\div$ {same determinant without first row and column}

$$= -\begin{vmatrix} 0 & F_{0,0} & F_{1,0} & F_{0,1} & F_{1,1} \\ F_{0,0} & (F_{0,0}+wF_{1,0}+zF_{0,1}) & (F_{1,0}+wF_{2,0}+zF_{1,1}) & (F_{0,1}+wF_{1,1}+zF_{0,2}) & (F_{1,1}+wF_{2,1}+zF_{1,2}) \\ F_{1,0} & (F_{1,0}+wF_{2,0}+zF_{1,1}) & (F_{2,0}+wF_{3,0}+zF_{2,1}) & (F_{1,1}+wF_{2,1}+zF_{1,2}) & (F_{2,1}+wF_{3,1}+zF_{2,2}) \\ F_{0,1} & (F_{0,1}+wF_{1,1}+zF_{0,2}) & (F_{1,1}+wF_{2,1}+zF_{1,2}) & (F_{0,2}+wF_{1,2}+zF_{0,3}) & (F_{1,2}+wF_{2,2}+zF_{1,3}) \\ F_{1,1} & (F_{1,1}+wF_{2,1}+zF_{1,2}) & (F_{2,1}+wF_{3,1}+zF_{2,2}) & (F_{1,2}+wF_{2,2}+zF_{1,3}) & (F_{2,2}+wF_{3,2}+zF_{2,3}) \end{vmatrix}$$

$\div$ {same determinant without first row and column}.

Possible linear dependences among the basis vectors can be taken care of directly in the final (numerical) form of the approximant by following the procedure described after (2.18). In a similar way we obtain for example

$$G(0,1,0) = \langle f,f\rangle + \begin{vmatrix} 0 & \langle f,(wA+zB)f\rangle & \langle Af,(wA+zB)f\rangle \\ \langle f,(wA+zB)f\rangle & \langle f,(wA+zB)(1+wA+zB)f\rangle & \langle Af,(wA+zB)(1+wA+zB)f\rangle \\ \langle Af,(wA+zB)f\rangle & \langle Af,(wA+zB)(1+wA+zB)f\rangle & \langle Af,(wA+zB)(1+wA+zB)Af\rangle \end{vmatrix}$$

(3.13) $\div$ {same determinant without first row and column}

$$= F_{0,0} + \begin{vmatrix} 0 & (wF_{1,0}+zF_{0,1}) & (wF_{2,0}+zF_{1,1}) \\ (wF_{1,0}+zF_{0,1}) & (wF_{1,0}+zF_{0,1}+w^2F_{2,0}+2wzF_{1,1}+z^2F_{0,2}), & (wF_{2,0}+zF_{1,1}+w^2F_{3,0}+2wzF_{2,1}+z^2F_{1,2}) \\ (wF_{2,0}+zF_{1,1}) & (wF_{2,0}+zF_{1,1}+w^2F_{3,0}+2wzF_{2,1}+z^2F_{1,2}) & (wF_{3,0}+zF_{2,1}+w^2F_{4,0}+2wzF_{3,1}+z^2F_{2,2}) \end{vmatrix}$$

$\div$ {same determinant without first row and column}.

We observe the following feature of the approximants $J(N, R, S)$ and $G(N, R, S)$. If we treat the assumed given coefficients, say (3.6), (3.7), and (3.8), as fixed, and ask what other variationally optimal bounds $J(\Phi_{\text{opt}})$ can be imposed on $F(w, z)$ using only these coefficients, then in fact any other suitable basis set will span a subspace of the space spanned by the vectors in (α) and (β). For example, in § 5 we consider the PA bounds on $F(w, z)$ and see that these derive from a much smaller basis set whilst utilizing comparable sets of coefficients. This means that the bounds $J(N, R, S)$ and $G(N, R, S)$ are the best that can be inferred from $J(\Phi)$ and $G(\Psi)$, assuming one knows only those coefficients $F_{m,n}$ which are actually used to construct the bounds.

It follows from the variationally optimal nature of the bounds

$$J(N, R, S) \leqq F(w, z) \leqq G(N, R, S), \qquad w \geqq 0, \quad z \geqq 0, \tag{3.14}$$

that they must improve as the basis set is enlarged. Thus we obtain, for allowed (N, R, S), $(N, R+1, S)$, and $(N, R, S+1)$:

$$\begin{aligned} J(N, R, S) &\leqq J(N, R+1, S); \qquad G(N, R+1, S) \leqq G(N, R, S), \\ J(N, R, S) &\leqq J(N, R, S+1); \qquad G(N, R, S+1) \leqq G(N, R, S) \end{aligned} \tag{3.15}$$

and

$$J(N, R, S) \leqq J(N+1, R, S); \qquad G(N+1, R, S) \leqq G(N, R, S). \tag{3.16}$$

When the set of vectors (β) is empty ($R+S=N$), $J(N, R, S)$ is actually identical with the two-variable approximant $F_{N+1}(w, z)$ suggested by Alabiso and Butera [4], on moment-theoretic grounds, as a possible generalization to two variables of the $[N/(N+1)]$ PA. When $F(w, z)$ belongs to a class of *extended* Stieltjes functions, they prove that the sequence $\{F_{N+1}(w, z): N=0, 1, \cdots\}$ converges to $F(w, z)$ for w and z lying in certain complex domains. Their theorem applies in particular to the functions (1.1) considered in this paper, when $w \geqq 0$ and $x \geqq 0$, provided that the two sums $\sum_{m=0}^{\infty} \{F_{2m,0}\}^{-1/(2m)}$ and $\sum_{n=0}^{\infty} \{F_{0,2n}\}^{-1/(2n)}$ are divergent. However, the approximants $J(N, R, S)$ with $R+S \neq N$, together with all of the $G(N, R, S)$ approximants appear to be new, as does the observation of their complementary bounding properties with respect to two-variable Stieltjes functions (1.1) and their power series matching properties, proved in § 4.

4. Matching property of the series expansions of *J(N, R, S)* and *G(N, R, S)* with the series expansion of *F(w, z)*. It was shown in [3] that the power series matching properties

$$[N/(N+1)] - f(\lambda) \sim \text{terms of order } \lambda^{2N+2} \text{ and higher,} \tag{4.1}$$

$$[N+1/(N+1)] - f(\lambda) \sim \text{terms of order } \lambda^{2N+3} \text{ and higher} \tag{4.2}$$

of the PA's for a single-variable Stieltjes function $f(\lambda)$, could be obtained from the variational characterization of these approximants. Here we apply parallel reasoning to $J(N, R, S)$, and $G(N, R, S)$, and establish the matching properties (1.6) and (1.7). In particular, when $R+S=N$ the double series expansion of each approximant agrees with the expansion (1.2) of $F(w, z)$ through precisely those terms which correspond to the coefficients used to construct the approximants.

We treat $J(N, R, S)$ first. It follows from (2.17) that Φ_{opt} is the unique solution of the problem

$$(1+PLP)\Phi_{\text{opt}} = Pf, \tag{4.3}$$

so that we can write

$$J(\Phi_{\text{opt}}) = \langle \Phi_{\text{opt}}, f\rangle = \langle \Phi_{\text{opt}}, Pf\rangle = \langle Pf, (1+PLP)^{-1}Pf\rangle. \tag{4.4}$$

We note here, incidentally, that it is in effect precisely this expression on the right-hand-side of (4.4) which is used by Alabiso and Butera [4] as their fundamental approximant to $F(w, z)$. Thus, the identification of their $F_N(w, z)$ with our $J(N, R, S)$ when $R+S=N$ is immediate, because their basis set is the same as ours in this case.

From (4.4) it follows that

$$J(N, R, S) \sim \sum_{k=0}^{\infty} (-1)^k \langle f, (PLP)^k f\rangle \tag{4.5}$$

where we understand $(PLP)^0 = P$, and where P now refers to the basis set of § 3, indexed by (N, R, S). Hence, the term of order $w^m z^n$ in the double-series expansion of $J(N, R, S)$ arises from $(-1)^{m+n}\langle f, (PLP)^{m+n} f\rangle$, $m \geqq 0$, $n \geqq 0$.

Consider

$$S_k = \langle f, (PLP)^k f\rangle \quad \text{where} \quad 0 \leqq k \leqq 2N+1; \tag{4.6}$$

we can always rewrite this as

$$S_k = \langle (PLP)^{m'} f, L^{\Theta}(PLP)^{n'} f\rangle \tag{4.7}$$

where $0 \leqq m' \leqq N$, $0 \leqq n' \leqq N$, $\Theta = 0$ or 1, and $m'+n'+\Theta = k$. Look at $(PLP)^r f$ for $0 \leqq r \leqq N$. We have

$$(PLP)^0 f = Pf = f \tag{4.8}$$

since the set (α) contains f. Again,

$$(PLP)^1 f = PLf = P(wAf + zBf) = wAf + zBf = Lf \tag{4.9}$$

because the set (α) contains Af and Bf. Similarly,

$$\begin{aligned} (PLP)^2 f &= PL(PLP)f = PL^2 f \\ &= P(w^2A^2f + 2wzABf + z^2B^2f) = L^2 f \end{aligned} \tag{4.10}$$

since the set (α) contains A^2f, ABf, and B^2f. Groing on in this style we eventually obtain

$$(PLP)^r f = L^r f \quad \text{for all } 0 \leqq r \leqq N, \tag{4.11}$$

because the set (α) contains all vectors $A^k B^l f$ with $k \geqq 0$, $l \geqq 0$, and $1 \leqq k+l \leqq N$, together with f. Hence we can rewrite (4.7) as

$$S_k = \langle L^{m'} f, L^{\Theta} L^{n'} f\rangle = \langle f, L^{m'+n'+\Theta} f\rangle = \langle f, L^k f\rangle, \tag{4.12}$$

for $0 \leqq k \leqq 2N+1$. It follows that the coefficient of $w^m z^n$ in the formal expansion of $J(N, R, S)$ is

$$(4.13) \qquad (-1)^{m+n}\binom{m+n}{n}\langle f, A^m B^n f\rangle = (-1)^{m+n}\binom{m+n}{n} F_{m,n}$$

for each $m \geqq 0$, $n \geqq 0$, such that $0 \leqq m+n \leqq 2N+1$. This establishes (1.6).

We now prove (1.7). We assume for simplicity that L is strictly positive, and afterwards indicate how the proof is modified when this is not the case. From (2.19) and (2.20) we have

$$(4.14) \qquad G(N, R, S) = \langle f, f\rangle - \langle PLf, [PL(1+L)P]^{-1} PLf\rangle$$

where $[PL(1+L)P]^{-1}$ is the *reduced* inverse, defined by

$$(4.15) \qquad [PL(1+L)P]^{-1}[PL(1+L)P] = [PL(1+L)P][PL(1+L)P]^{-1} = P.$$

We introduce the notations

$$(4.16) \qquad \bar{L} = PLP, \quad \text{and} \quad \overline{L^2} = PL^2P,$$

and denote the reduced inverse of $\bar{L}$ by $\bar{L}^{-1}$ so that

$$(4.17) \qquad \bar{L}\bar{L}^{-1} = \bar{L}^{-1}\bar{L} = P.$$

Then we have the formal expansion

$$(4.18) \qquad [PL(1+L)P]^{-1} \sim \sum_{k=0}^{\infty} (\bar{L}^{-1}\overline{L^2})^k \bar{L}^{-1}(-1)^k$$

wherein we take

$$(4.19) \qquad (\bar{L}^{-1}\overline{L^2})^0 = P.$$

Since the basis set (α) contains f we can rewrite (4.14) as

$$(4.20) \qquad G(N, R, S) = \langle f, f\rangle - \langle \bar{L}f, [PL(1+L)P]^{-1}\bar{L}f\rangle,$$

and substituting from (4.18) we derive

$$(4.21) \qquad G(N, R, S) \sim \langle f, f\rangle - \sum_{k=0}^{\infty} \langle \bar{L}f, (\bar{L}^{-1}\overline{L^2})^n f\rangle.$$

By looking at the terms in this expression when $w = \lambda z$ we deduce that the term in $w^m z^n$, $m+n \geqq 1$, can only arise from $\langle Lf, (\bar{L}^{-1}\overline{L^2})^{m+n-1} f\rangle$, whilst the constant term is simply $\langle f, f\rangle = F_{0,0}$. Accordingly, we look at

$$(4.22) \qquad T_k = \langle \bar{L}f, (\bar{L}^{-1}\overline{L^2})^k f\rangle \quad \text{for } 0 \leqq k \leqq 2N+1.$$

We have

$$T_0 = \langle PLPf, Pf\rangle = \langle f, Lf\rangle \tag{4.23}$$

because (α) contains f. For $1 \leqq k \leqq 2N+1$, we can always rewrite (4.22) as

$$T_k = \langle (\bar{L}^{-1}\overline{L^2})^{m'} f, L^2(\bar{L}^{-1}\overline{L^2})^{n'} f\rangle \tag{4.24}$$

where $0 \leqq m' \leqq N$ and $0 \leqq n' \leqq N$. We now claim that

$$(\bar{L}^{-1}\overline{L^2})^r f = \begin{cases} L^r f & \text{when } r = 1, 2, \cdots, N, \\ f & \text{when } r = 0. \end{cases} \tag{4.25}$$

For $r = 0$ we have

$$(\bar{L}^{-1}\overline{L^2})^0 f = Pf = f \tag{4.26}$$

since (α) contains f. For $r = 1$ we have

$$\begin{aligned}(\bar{L}^{-1}\overline{L^2})^1 f &= \bar{L}^{-1}P\overline{L^2}Pf = \bar{L}^{-1}PLPLf \\ &= (\bar{L}^{-1}\bar{L})Lf = PLf = Lf\end{aligned} \tag{4.27}$$

where we have used the fact that (α) contains f, Af, and Bf to give the second and fifth equalities, and (4.17) to give the fourth equality. For $r = 2$ we have

$$\begin{aligned}(\bar{L}^{-1}\overline{L^2})^2 f &= \bar{L}^{-1}\overline{L^2}Lf = \bar{L}^{-1}PL^2PLf \\ &= \bar{L}^{-1}PLPL^2f = PL^2f = L^2 f\end{aligned} \tag{4.28}$$

where we have used (4.27), and made repeated use of the fact that (α) contains A^2f, ABf, and B^2f. Going on in this fashion, and bearing in mind the contents of (α), we verify (4.25). Substituting (4.25) in (4.24) now yields

$$T_k = \langle f, L^k f\rangle \quad \text{for } k = 1, 2, \cdots, 2N+1; \tag{4.29}$$

from which we infer (1.7), provided L is strictly positive.

When L is not strictly positive, the projection operator P occurring on the far right-hand-sides of (4.15), (4.17), and (4.19), must be replaced by P_L, the projection operator associated with the subspaces of $\mathscr{h}$ spanned by the eigenvectors of PLP with nonzero eigenvalues. All of the above arguments still go through, but one must use the relations

$$P_L L^n f = L^n f \quad \text{for } n = 1, 2, \cdots, N \tag{4.30}$$

and

$$LP_L f = Lf, \tag{4.31}$$

which follow from the definition of P_L.

Although the above proofs refer explicitly to function $F(w, z)$ of the form (1.1), it seems very likely that the matching properties (1.6) and (1.7) are true for an arbitrary function $F(w, z)$ provided that the determinants in the expressions which define the approximants, for example (3.12) and (3.13), are nonsingular (i.e. provided that one does not have to omit various rows and columns to allow for linear dependences). We conjecture that the formal power series matching property of these expressions could be proved algebraically: one piece of evidence for this is similarity between the formulas for $J(N, R, S)$ and $G(N, R, S)$ and Nuttall's compact formulas for one variable PA's: the power series matching property of the latter *is* algebraic.

5. Comparison of $J(N, R, S)$ and $G(N, R, S)$ with Padé approximants. Given an initial set of coefficients in the double series expansion (1.2), a natural way of imposing upper and lower bounds on $F(w, z)$, $w \geqq 0$, $z \geqq 0$, is to construct Padé approximants [5].

For the series

$$f(\lambda) \sim \sum_{n=0}^{\infty} f_n(-\lambda)^n \tag{5.1}$$

the $[M/N]$ PA is defined according to

$$[M/N] = \sum_{r=0}^{M} p_r\lambda^r \Big/ \sum_{s=0}^{N} q_s\lambda^s \tag{5.2}$$

where the coefficients, the p_r's and q_r's, are specified by

$$\left(\sum_{s=0}^{N} q_s\lambda^s\right)\left(\sum_{n=0}^{N+M} f_n(-\lambda)^n\right) - \sum_{r=0}^{M} p_r\lambda^r = \text{terms of order } \lambda^{N+M+1} \text{ and higher} \tag{5.3}$$

together with the relative normalization condition

$$p_0 = 1. \tag{5.4}$$

When $f(\lambda)$ is a Stieltjes function in λ, the various $[M/N]$ PA's are guaranteed to exist and they impose, among others, the bounds

$$[N/(N+1)] \leqq f(\lambda) \leqq [N+1/(N+1)] \quad \text{when } \lambda \geqq 0. \tag{5.5}$$

Now, for each fixed positive w and z, the function

$$F(w, z; \lambda) = \int_0^{\infty}\int_0^{\infty} \frac{d\sigma(s, t)}{(1+\lambda(wz+zt))} \tag{5.6}$$

in a Stieltjes function in the single variable λ, with formal series

$$F(w, z; \lambda) \sim \sum_{n=0}^{\infty} F_n(w, z)(-\lambda)^n \tag{5.7}$$

where

$$F_n(w, z) = \sum_{\substack{p+q=n \\ p\geqq 0,\, q\geqq 0}} \binom{n}{p} F_{p,q} w^p z^q. \tag{5.8}$$

Hence, if we are given the sets of coefficients (3.6) and (3.9), we can correspondingly construct the lower and upper bounds in (5.5) when $\lambda = 1$, which we denote respectively by $[N/(N+1)](w, z)$ and $[N+1/(N+1)](w, z)$. Then we have

$$[N/(N+1)](w, z) \leqq F(w, z) \leqq [N+1/(N+1)](w, z), \qquad w \geqq 0, \quad z \geqq 0. \tag{5.9}$$

These approximants have been suggested by Graves-Morris (see [13]), although their bounding properties were not noted.

Here we compare $J(N, R, S)$ with $[N/(N+1)](w, z)$ and $G(N, R, S)$ with $[N+1/(N+1)](w, z)$. We concentrate first on the case $R+S=N$, writing for brevity

$$J(N, R, S) = J(N) \quad \text{and} \quad G(N, R, S) = G(N) \quad \text{when } R+S=N. \tag{5.10}$$

$J(N)$ and $[N/(N+1)](w, z)$ share the following features: they both impose a lower bound on $F(w, z)$ for positive w and z; they both require exactly the same set of coefficients, namely (3.6), for their construction; and, with the use of (4.1), $[N/(N+1)]$ has the power series matching property

$$[N/(N+1)](w,z)-F(w,z)\sim \text{terms of order } w^p z^q, \qquad p+q \geqq 2N+2, \tag{5.11}$$

in common with $J(N)$ [cf. (1.6)].

Similarly, $G(N)$ and $[N+1/(N+1)](w,z)$ share the following features: they both impose an upper bound on $F(w,z)$ for positive w and z; they both require exactly the same set of coefficients, namely (3.9), for their construction; and, with the use of (4.2), $[N+1/(N+1)]$ has the property

$$[N+1/(N+1)](w,z)-F(w,z)\sim \text{terms of order } w^p z^q, \qquad p+q \geqq 2N+3, \tag{5.12}$$

in common with $G(N)$ [cf. (1.7)].

In order to see the structures of $[N/(N+1)](w,z)$ and $[N+1/(N+1)](w,z)$ we observe that they can be obtained from the optimized dual variational bounds $J(\Phi_{\text{opt}})$ and $G(\Psi_{\text{opt}})$, respectively, when the basis set is chosen to be (see [3]; or believe, without proof, the introductory remarks in § 3):

$$\Theta_1 = f; \qquad \Theta_n = L^{n-1}f, \qquad n=2,3,\cdots,N, \tag{5.13}$$

[The proofs of (1.6) and (1.7) given in § 4 apply equally well in the case of the basis set (5.13), thereby establishing (5.10) and (5.11)—if we hadn't known these already!]. Substituting the basis set (5.13) into (2.18) and (2.20), and allowing for the explicit (w,z) dependences of these vectors, in distinction to the case where (2.21) and (2.22) apply, we find that we can write

$$[N/(N+1)](w,z)=\sum_{\substack{n+m=N(N+1)\\ n\geqq 0, m\geqq 0}}^{n+m=N(N+2)} \alpha_{m,n} w^m z^n \Bigg/ \sum_{\substack{n+m=N(N+1)\\ n\geqq 0, m\geqq 0}}^{n+m=N(N+2)+1} \beta_{m,n} w^m z^n \tag{5.14}$$

and

$$[N+1/(N+1)](w,z)=\sum_{\substack{n+m=(N+1)^2\\ n\geqq 0, m\geqq 0}}^{n+m=(N+1)(N+2)} \gamma_{m,n} w^m z^n \Bigg/ \sum_{\substack{n+m=(N+1)^2\\ n\geqq 0, m\geqq 0}}^{n+m=(N+1)(N+2)} \rho_{m,n} w^m z^n \tag{5.15}$$

where the $\alpha_{m,n}$'s $\beta_{m,n}$'s $\gamma_{m,n}$'s, and $\rho_{m,n}$'s, are real constants. When $w=\lambda z$ with λ a constant, $[N/(N+1)](w,z)$ reduces to a polynomial of degree N in z divided by one degree $(N+1)$, while $[N+1/N+1](w,z)$ becomes a ratio of two polynomials of degree $(N+1)$ in z. On the other hand $J(N)$ and $G(N)$ can be expressed in the forms (2.21) and (2.22) with $\mathcal{N}=\frac{1}{2}(N+1)(N+2)$. *Thus, viewed purely as rational expressions,* $J(N)$ *and* $G(N)$ *are more complicated than* $[N/(N+1)](w,z)$ *and* $[N+1/(N+1)](w,z)$ because they involve a greater number of "unknown constants".

Despite this difference, we note that in the case $w=\lambda z$ with λ a positive constant $[N/(N+1)](w,z)$ and $[N+1/(N+1)](w,z)$ can be decomposed respectively into the forms (2.26) and (2.27) with $\mathcal{N}=N$. To prove this we observe that when $w=\lambda z$, $\lambda>0$, $z>0$, the basis set (5.13) can, without loss of generality, be replaced by the set

$$\Theta_1 = f, \qquad \Theta_n=(A+\lambda B)^{n-1} \quad \text{for } n=2,3,\cdots,N, \tag{$\widetilde{5.13}$}$$

which *is* (w,z)-independent, so that the derivation of (2.26) and (2.27) pertains here with $\mathcal{N}=N$. Similarly, we also know that in this case $J(N)$ and $G(N)$ have the structures (2.26) and (2.27) respectively, but with $\mathcal{N}=\frac{1}{2}(N+1)(N+2)$. Hence, the *qualitative* behavior of the pair $[N/(N+1)](w,z)$ and $[N+1/(N+1)](w,z)$ is essentially the same as that of $J(N)$ and $G(N)$.

Although they are constructed using identical sets of coefficients, we have the *quantitative* relationships

$$[N/(N+1)](w,z) \leqq J(N); \quad \text{and} \quad G(N) \leqq [N+1/(N+1)](w,z), \qquad w \geqq 0, \quad z \geqq 0, \tag{5.16}$$

because the basis set (5.13) spans a subspace of the space spanned by (α). In general, the subspace spanned by (α) and (β) has dimension $\mathcal{N}(N, R, S) = \frac{1}{2}(N+1)(N+2)+(R+S-N)$, while (5.13) spans a subspace of dimension N, so one can expect that the bounds (3.14) are a dramatic improvement Over (5.9) for large N.

When $R+S>N$, the approximants $J(N, R, S)$ and $G(N, R, S)$ use sets of $F_{m,n}$'s which contain those used by $[N/(N+1)](w, z)$ and $[N+1/(N+1)](w, z)$, but these larger sets cannot be used to derive PA's of higher order. They provide even better bounds than $J(N)$ and $G(N)$ (recall (3.15)), but otherwise their qualitative features are similar to those of $J(N)$ and $G(N)$.

The approximants $[N/(N+1)](w, z)$ and $[N+1/(N+1)](w, z)$ have the advantages that they can be simply constructed with the aid of the ε-algorithm [13], and that their bounding properties apply not only to $F(w, z)$ in (1.1), but in fact to any function expressible in the form (2.1) when the operators A and B do *not* commute, $w \geqq 0$, $z \geqq 0$. $J(N, R, S)$ and $G(N, R, S)$ have the advantages that they impose tighter bounds than the PA's using the same sets of given $F_{m,n}$'s, and they can be constructed from a wider variety of sets of given $F_{m,n}$'s.

6. Applications and examples. The most obvious application of the preceding theory is to the evaluation of upper and lower bounds on special mathematical functions which can be written in the form (1.1), and in particular to the numerical evaluation of various double integrals of this structure. Among the special functions which can be expressed in the form (1.1) are a number of the two-variable hypergeometric functions; see for example [15].

Example 1. We take

$$F(w,z)=\frac{1}{w}\int_0^\infty \exp(-t^2)\ln\left(\frac{1+zt+2w}{1+zt+w}\right)dt=\int_1^2 ds\int_0^\infty \frac{dt\exp(-t^2)}{(1+ws+zt)}, \tag{6.1}$$

which has the form (1.1), and for which the coefficients

$$F_{m,n}=\frac{(2^{m+1}-1)}{(m+1)}\int_0^\infty t^n\exp(-t^2)\,dt, \qquad m, n=0, 1, \cdots, \tag{6.2}$$

are available. For the purpose of comparing the $[N/(N+1)](w, z)$ and $[N+1/(N+1)](w, z)$ approximants with $J(N, R, S)$ and $G(N, R, S)$ we suppose that in fact we know only the coefficients

$$\{F_{0,0}; F_{1,0}; F_{0,1}; F_{2,0}; F_{1,1}; F_{0,2}; F_{3,0}; F_{2,1}; F_{0,3}; F_{4,0}; F_{3,1}; F_{2,2}\}. \tag{6.3}$$

Using this set of information the best PA bounds we can construct are

$$[1/2](w,z) \leqq F(w,z) \leqq [1/1](w,z), \qquad w \geqq 0, \quad z \geqq 0; \tag{6.4}$$

and the best J, G bounds are

$$J(1,1,0) \leqq F(w,z) \leqq G(0,1,0), \qquad w \geqq 0, \quad z \geqq 0. \tag{6.5}$$

We find that, at $w = z = 1$,

$$[1/2](1, 1) = 0.296874 < J(1, 1, 0) = 0.296902 \leqq F(1, 1)$$

and

$$[1/1](1, 1) = 0.313229 > G(0, 0, 1) = 0.308936 \geqq F(1, 1).$$

Thus, there is some advantage in using the J, G bounds here which is in accord with the discussion in § 5. However, the PA bounds were easier to evaluate.

One is sometimes faced with the following situation: it is given that a function $\mathscr{F}(z)$ is a Stieltjes function, say

$$\mathscr{F}(z) = \int_0^\infty \frac{d\tilde{\sigma}(t)}{(1+tz)}; \tag{6.6}$$

where $\tilde{\sigma}(t)$ is bounded and nondecreasing on $0 \leqq t < \infty$; and, in addition to knowing the first few "integer" moments from the sequence

$$\mathscr{F}_n = \int_0^\infty t^n \, d\tilde{\sigma}(t), \qquad n = 0, 1, \cdots, \tag{6.7}$$

from which bounds on $\mathscr{F}(z)$, $z \geqq 0$, can be inferred by using the usual one-variable PA's, one also knows a corresponding set of "half-integer" moments from the sequence

$$\mathscr{F}_{n+1/2} = \int_0^\infty t^{n+1/2} \, d\tilde{\sigma}(t), \qquad n = 0, 1, \cdots. \tag{6.8}$$

One would like to be able to use the latter additional information to improve upon the usual PA bounds on $\mathscr{F}(z)$. This situation arises, for example, in the case of the quantum-mechanical dynamic polarizability function associated with a ground-state atom or molecule [15], where one knows various sum-rules of both even and odd orders.

We show how J, G approximants can be used in such a case to improve upon the one-variable PA bounds. We set

$$F(w, z) = \int_0^\infty \int_0^\infty \frac{d\sigma(s, t)}{(1 + ws + zt)}, \tag{6.9}$$

with

$$d\sigma(s, t) = \delta(s - t^{1/2}) \, d\tilde{\sigma}(t) \, ds, \tag{6.10}$$

$\delta(r)$ denoting the delta-function in the variable r, so that

$$F(0, z) = \mathscr{F}(z). \tag{6.11}$$

Then we find correspondingly

$$F_{m,n} = \int_0^\infty s^m t^n \delta(s - t^{1/2}) \, d\tilde{\sigma}(t) \, ds = \int_0^\infty t^{(n+m/2)} \, d\tilde{\sigma}(t), \tag{6.12}$$

whence, for each m and n, $F_{m,n}$ is either an "integer" or "half-integer" moment of the distribution $\tilde{\sigma}(t)$. Bounds on $\mathscr{F}(z)$ when $0 \leqq z < \infty$ which utilize "half-integer" moments as well as the usual Taylor series coefficients (needed for one-variable PA bounds) now follow from $J(N, R, S)$ and $G(N, R, S)$, wherein we set $w = 0$. Some of the basis vectors in (α) and (β) may now be linearly dependent (in fact equal) and this must be taken into account. This can be achieved either by noting linear dependences among rows in the determinantal forms describing the approximants, or else, more simply, by noting that here

$$A^m B^n f = B^{n+m/2} f, \tag{6.13}$$

which reduces the number of basis vectors which need to be included when writing down the approximants. A typical approximant looks like

$$J(1,1,1) = -\begin{vmatrix} 0 & \mathscr{F}_0 & \mathscr{F}_{1/2} & \mathscr{F}_1 & \mathscr{F}_{3/2} \\ \mathscr{F}_0 & (\mathscr{F}_0+z\mathscr{F}_1) & (\mathscr{F}_{1/2}+z\mathscr{F}_{3/2}) & (\mathscr{F}_1+z\mathscr{F}_2) & (\mathscr{F}_{3/2}+z\mathscr{F}_{5/2}) \\ \mathscr{F}_{1/2} & (\mathscr{F}_{1/2}+z\mathscr{F}_{3/2}) & (\mathscr{F}_1+z\mathscr{F}_2) & (\mathscr{F}_{3/2}+z\mathscr{F}_{5/2}) & (\mathscr{F}_2+z\mathscr{F}_3) \\ \mathscr{F}_1 & (\mathscr{F}_1+z\mathscr{F}_2) & (\mathscr{F}_{3/2}+z\mathscr{F}_{5/2}) & (\mathscr{F}_2+z\mathscr{F}_3) & (\mathscr{F}_{5/2}+z\mathscr{F}_{7/2}) \\ \mathscr{F}_{3/2} & (\mathscr{F}_{3/2}+z\mathscr{F}_{5/2} & (\mathscr{F}_2+z\mathscr{F}_3) & (\mathscr{F}_{5/2})+z\mathscr{F}_{7/2}) & (\mathscr{F}_3+z\mathscr{F}_4) \end{vmatrix} \tag{6.14}$$

$\div$ {same determinant without first row and column}.

Example 2. Consider the hypergeometric function

$$F(1,1;2;-z) = \mathscr{F}(z) = \int_0^1 \frac{dt}{(1+tz)} = \frac{\ln(1+z)}{z} \tag{6.15}$$

for which

$$\mathscr{F}_p = \int_0^1 t^p\,dt = 1/(p+1), \qquad p = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \cdots. \tag{6.16}$$

Using $\mathscr{F}_0$, $\mathscr{F}_{1/2}$, $\mathscr{F}_1$, $\mathscr{F}_{1\,1/2}$, and $\mathscr{F}_2$, we find $J(0, 1, 0)$ is a polynomial of degree one in z divided by a polynomial of degree two in z. At $z = 1$ we get the lower bound

$$J(0,1,0)_{z=1,(w=0)} = 0.693069 \leqq \ln 2 = 0.693147.$$

A comparable PA is $[1/2]$, which uses $\mathscr{F}_0$, $\mathscr{F}_1$, $\mathscr{F}_2$, $\mathscr{F}_3$, and yields

$$[1/2]_{z=1} = 0.69203 \leqq \ln 2,$$

this bound being significantly worse than $J(0, 1, 0)_{z=1,(w=0)}$ despite the fact that the PA is a similar ratio of polynomials. However, for small enough positive z the PA bound would be the best because its Taylor series expansion agrees with $(\ln(1+z))/z$ through order z^3 compared with z^2 for $J(0, 1, 0)$.

The rational approximants $J(N, R, S)$ and $G(N, R, S)$ can clearly be used to derive bounds on quantities represented by double perturbation series of the form

$$F(w,z) \sim \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} (-1)^{m+n} \binom{n+m}{n} \langle f, A^m B^n f\rangle w^m z^n \tag{6.17}$$

when the operators A and B are self-adjoint, positive, and commute. In such a case they would be more efficient than the approximants $[N/(N+1)](w, z)$ and $[N+1/(N+1)]$ (w, z), as described in § 5. However, can they be used to sum double-perturbation series associated with quantities of the form

$$\tilde{F}(w,z) = \langle f, (1 + w\tilde{A} + z\tilde{B})^{-1} f\rangle \tag{6.18}$$

when the linear operators $\tilde{A}$ and $\tilde{B}$ are self-adjoint, positive, but do *not* commute?

Series of the latter type can occur in below threshold scattering theory when there are two coupling constants w and z, and (6.18) is one way of writing the forward scattering amplitude (see [14], for example). In this situation the two-variable PA's of § 5 can still be used to impose upper and lower bounds on $\tilde{F}(w, z)$. It is suggested by Alabiso and Butera [4] that the approximants $\tilde{F}_{N+1}(w, z)$ might provide good estimates for $\tilde{F}(w, z)$: we use the following simple example to argue that in general the approximants $J(N, R, S)$ cannot be used to advantage in the "noncommuting" case.

Example 3. LeT $\phi(x)$ be the unique solution of the problem

$$\left(1-w\frac{d^2}{dx^2}+zx^2\right)\phi(x)=e^{-x^2}, \qquad -\infty<x<\infty, \quad w\geqq 0, \quad z\geqq 0, \tag{6.19}$$

subject to the condition that $\phi(x)$ is square-integrable over the real line. Let $\mathscr{h}$ be the real Hilbert space which is associated with the inner product

$$\langle f_1, f_2\rangle=\int_{-\infty}^{+\infty} f_1(x)f_2(x)\,dx. \tag{6.20}$$

Then (6.18) takes the form (2.4) with $A(=\tilde{A})$ and $B(=\tilde{B})$ respectively equal to $-d^2/dx^2$ and "multiplication by x^2", complete with

$$D(\tilde{A})=\{\psi(x)\in \mathscr{h} : \psi''(x) \text{ exists}, \psi''\in \mathscr{h}\} \tag{6.21}$$

and

$$D(\tilde{B})=\{\eta(x)\in \mathscr{h} : x^2\eta(x)\in \mathscr{h}\}. \tag{6.22}$$

Both operators are linear, self-adjoint, and positive, but they do not commute. We look at the quantity

$$\tilde{F}(w, z)=\left\langle \phi(x), \left(\frac{2}{\pi}\right)^{1/4} e^{-x^2}\right\rangle=\left(\frac{2}{\pi}\right)^{1/4}\int_{-\infty}^{+\infty}\phi(x)\,e^{-x^2}\,dx. \tag{6.23}$$

In the formal expansion of this function in ascending powers of w and z, the coefficient of $w^m z^n$ arises from

$$(-1)^{m+n}\left(\frac{2}{\pi}\right)^{1/2}\int_{-\infty}^{+\infty} e^{-x^2}\left(-w\frac{d^2}{dx^2}+zx^2\right)^{m+n} e^{-x^2}\,dx, \tag{6.24}$$

and is finite for all m and n. Writing the double expansion in the form (1.2) we can verify that

$$\tilde{F}_{0,0}=1;\quad \tilde{F}_{1,0}=1;\quad \tilde{F}_{0,1}=1/4;\quad \tilde{F}_{2,0}=3;\quad \tilde{F}_{1,1}=-1/4;\quad \tilde{F}_{3,0}=15; \quad \tilde{F}_{2,1}=-11/12. \tag{6.25}$$

At $w=z=1$, the best PA bounds we can impose on $\tilde{F}(1, 1)$, on the basis of (6.24), are

$$[0/1](1,1)=4/9=0.44\leqq \tilde{F}(1,1)\leqq 1=[0/0](1,1). \tag{6.26}$$

On the other hand, making fuller use of the information (6.24), we obtain the *estimate*

$$J(0,1,0)_{w=z=1}=0.48566\simeq \tilde{F}(1,1) \tag{6.27}$$

which, while it is sensible, is not as useful a result as (6.26). If we now incorporate the additional coefficients

$$\tilde{F}_{0,2}=3/16;\quad \tilde{F}_{1,2}=-11/48;\quad \tilde{F}_{0,3}=15/16 \tag{6.28}$$

into the set (6.25), we can obtain the PA bounds

$$[1/2](1,1)=0.47736 \leqq \tilde{F}(1,1) \leqq [1/1](1,1)=0.60317 \tag{6.29}$$

and the estimate

$$J(1,1,0)_{w=z=1}=0.44582 \simeq \tilde{F}(1,1). \tag{6.30}$$

The approximant $J(1,1,0)$ here is identical to the Alabiso–Butera approximant $\tilde{F}_2(w,z)$. The estimate (6.30) lies outside the rigorous bounds (6.29), and is a worse approximation than (6.27). Hence, in the "noncommuting" case, the approximants $J(N,R,S)$ do not in general provide reliable estimates; the PA's of § 5 provide bounds which are both rigorous and more accurate.

REFERENCES

[1] S. T. Epstein, *Variational upper and lower bounds on the dynamical polarizability at imaginary frequency*, J. Chem. Phys., 48 (1968), pp. 4716–4717.

[2] S. T. Epstein and M. F. Barnsley, *A variational approach to the theory of multipoint Padé approximants*, J. Math. Phys., 14 (1973), pp. 314–325.

[3] M. F. Barnsley and P. D. Robinson, *Dual variational principles and Padé-type approximants*, J. Inst. Math. Appl., 14 (1974), pp. 229–249.

[4] C. Alabiso and P. Butera, *N-variable rational approximants and method of moments*, J. Math. Phys., 16 (1975), pp. 840–845.

[5] G. A. Baker, Jr., *Essentials of Padé Approximants*, Academic Press, New York and London, 1975.

[6] J. S. R. Chisholm, *Rational approximants defined from double power series*, Math. Comp., 27 (1973), pp. 841–848.

[7] P. R. Graves-Morris, R. Hughes Jones and G. J. Makinson, *The calculation of some rational approximants in two variables*, J. Inst. Math. Appl., 13 (1974), pp. 311–320.

[8] A. K. Common and P. R. Graves-Morris, *Some properties of Chisholm approximants*, Ibid., 13, (1974), pp. 229–232.

[9] J. S. R. Chisholm and J. McEwan, *Rational approximants defined from power series in N-variables*, Proc. Roy. Soc. London Ser. A, 336 (1974), pp. 421–452.

[10] J. S. R. Chisholm and P. R. Graves-Morris, *Generalizations of the theorem of de Montessus to two-variable approximants*, Ibid., 342 (1975), pp. 341–372.

[11] B. Noble and M. J. Sewell, *On dual extremum principles in applied mathematics*, J. Inst. Math. Appl., 9 (1972), pp. 123–193.

[12] S. T. Epstein, *The Variational Method in Quantum Chemistry*, Academic Press, New York, 1974.

[13] P. J. S. Watson, *Two-variable rational approximants: a new method*, J. Phys. A., 50 (1974), pp. 167–170.

[14] G. A. Baker, *The Padé approximant method and some related generalizations*, The Padé Approximant in Theoretical Physics, G. A. Baker, Jr. and J. L. Gammel, eds., Academic Press, New York and London, 1970, pp. 33–34.

[15] A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, *Higher Transcendental Functions*, vol. 1, McGraw-Hill, New York, 1953, p. 230.

[16] P. W. Langhoff and M. Karplus, *Application of Padé approximants to dispersion force and optical polarizability computations*, The Padé Approximant in Theoretical Physics, G. A. Baker, Jr. and J. L. Gammel, eds., Academic Press, New York and London, 1970, pp. 41–97.

SIAM J. MATH. ANAL.
Vol. 9, No. 2, April 1978
Copyright © 1978 Society for Industrial and Applied Mathematics

A NOTE ON THE EIGENVALUES OF HERMITIAN MATRICES*

DAVID SLEPIAN† AND HENRY J. LANDAU‡

Abstract. Two simple relations are derived that connect the eigenvalues of a Hermitian matrix with those of the submatrix obtained by deleting a row and the corresponding column. The relations, which readily establish the interlacing of these two sets of eigenvalues, are used to obtain an upper bound for the largest eigenvalue and a lower bound for the smallest eigenvalue of a Hermitian matrix.

1. Notation and summary of results. In this note we point out two simple relations that hold between the eigenvalues of a Hermitian matrix and those of the Hermitian submatrix obtained by deleting a given row and the corresponding column. The relations, which readily establish the interlacing of these two sets of eigenvalues, are used to obtain an upper bound for the largest eigenvalue and a lower bound for the smallest eigenvalue of the matrix.

First some matters of notation. The scalar product of two k-dimensional column vectors $\mathbf{c}=(c_1, c_2, \cdots, c_k)^T$ and $\mathbf{d}=(d_1, d_2, \cdots, d_k)^T$ is defined by $(\mathbf{c}, \mathbf{d})= c_1 d_1^* + c_2 d_2^* + \cdots + c_k d_k^*$ where the asterisk denotes complex conjugate and T denotes transpose. If $(\mathbf{x}, \mathbf{y})=0$, $\mathbf{x}$ and $\mathbf{y}$ are said to be orthogonal. A number λ is said to be an eigenvalue of multiplicity k of the Hermitian matrix A if there exist exactly k linearly independent vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$ such that $A\mathbf{x}_i = \lambda \mathbf{x}_i$, $i=1, 2, \cdots, k$. While we shall always use the unqualified word "eigenvalue" to mean an eigenvalue of multiplicity $k \geqq 1$, it will simplify our exposition to speak sometimes of eigenvalues of multiplicity zero; the statement "λ is an eigenvalue of A of multiplicity zero" means that λ is not an eigenvalue of A.

In all that follows $n \geqq 2$ is an integer. Denote by A the $n \times n$ Hermitian matrix having elements $a_{ij} = a_{ji}^*$, $i, j = 1, 2, \cdots, n$. The following facts are well known [1]. The eigenvalues of A are real and are identical with the roots $\lambda_1, \lambda_2, \cdots, \lambda_n$ of the nth degree polynomial

$$P_A(\lambda) \equiv \det (A - \lambda I), \tag{1}$$

where I is the $(n \times n)$ unit matrix. If λ_j is a root of $P_A(\lambda)$ of multiplicity k, then λ_j is an eigenvalue of multiplicity k of A. We suppose the eigenvalues are labeled so that

$$\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_n. \tag{2}$$

Orthonormal eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ corresponding to these eigenvalues can be found. They satisfy the relations

$$A\mathbf{x}_j = \lambda_j \mathbf{x}_j, \tag{3}$$

$$(\mathbf{x}_i, \mathbf{x}_j) = \delta_{ij}, \qquad i, j = 1, 2, \cdots, n. \tag{4}$$

We will denote by α_j the last component of $\mathbf{x}_j$.

Let B represent the $((n-1)\times(n-1))$ matrix obtained from A by deleting the last row and column, and let μ_j and $\mathbf{y}_j$ $j=1, \cdots, n-1$, be the eigenvalues and orthonormal eigenvectors of B, respectively, with $\mu_1 \geqq \mu_2 \geqq \cdots \geqq \mu_{n-1}$.

* Received by the editors November 26, 1975, and in revised form August 13, 1976.
† Bell Laboratories, Murray Hill, New Jersey 07974 and University of Hawaii, Honolulu, Hawaii 96822.
‡ Bell Laboratories, Murray Hill, New Jersey 07974.

As a final notational matter, we write

$$\mathbf{a} \equiv \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{n-1n} \end{pmatrix} \tag{5}$$

for the $(n-1)$-dimensional column vector obtained from the last column of A by omitting the diagonal element a_{nn} and adopt the abbreviation

$$\beta_j \equiv (\mathbf{a}, \mathbf{y}_j), \qquad j = 1, 2, \cdots, n-1, \tag{6}$$

for the components of $\mathbf{a}$ in the coordinate system formed by the eigenvectors $\{\mathbf{y}_j\}$.

In § 2 we shall prove the following.

THEOREM 1. *All the solutions of*

$$\phi_n(\lambda) \equiv \lambda - a_{nn} - \sum_{j=1}^{n-1} \frac{|\beta_j|^2}{\lambda - \mu_j} = 0 \tag{7}$$

considered as an equation for λ, *are eigenvalues of* A. *Among these solutions are all the eigenvalues of* A *that are not also eigenvalues of* B.

THEOREM 2. *All the solutions of*

$$\psi_n(\lambda) \equiv \sum_{j=1}^{n} \frac{|\alpha_j|^2}{\lambda - \lambda_j} = 0 \tag{8}$$

are eigenvalues of B. *Among these solutions are all the eigenvalues of* B *that are not also eigenvalues of* A.

Equations (7) and (8) are the relations referred to that connect the eigenvalues of A and those of B. In addition, we shall establish the identities

$$\psi_n(\lambda) = \frac{1}{\phi_n(\lambda)} = -\frac{\det(B - \lambda I)}{\det(A - \lambda I)}. \tag{9}$$

Theorem 2 and equation (9) are believed to be new: Theorem 1 is contained, for example, in material to be found in [1, pp. 94–98].

It follows immediately from Theorem 1 that if the eigenvalues of B are all of multiplicity one, so that

$$\mu_1 > \mu_2 > \cdots > \mu_{n-1} \tag{10}$$

and if all $n-1$ of the quantities $|\beta_j|$, $j = 1, 2, \cdots, n-1$, are different from zero, then ϕ_n in (7) has $n-1$ distinct poles. It is then clear from Fig. 1, which shows plots of $f_n(\lambda) \equiv \lambda - a_{nn}$ and $g_n(\lambda) \equiv f_n(\lambda) - \phi_n(\lambda)$, that the eigenvalues of B separate those of A, i.e. that

$$\lambda_1 > \mu_1 > \lambda_2 > \mu_2 > \cdots > \mu_{n-1} > \lambda_n. \tag{11}$$

These are the well-known "interlacing inequalities", [2, Thm. 4, p. 117].

Equation (7) will fail to have n real roots only when $\phi_n(\lambda)$ has fewer than $n-1$ distinct poles. This can occur because some of the eigenvalues of B are of multiplicity greater than one, or may arise because some of the quantities β_j are zero. A then inherits some of the eigenvalues of B. The situation can be summarized as follows. Let μ_j be an eigenvalue of B of multiplicity $k \geqq 1$. If $\mathbf{a}$ is orthogonal to all the eigenvectors of B belonging to μ_j and if μ_j is a root of (7), then μ_j is an eigenvalue of A of multiplicity

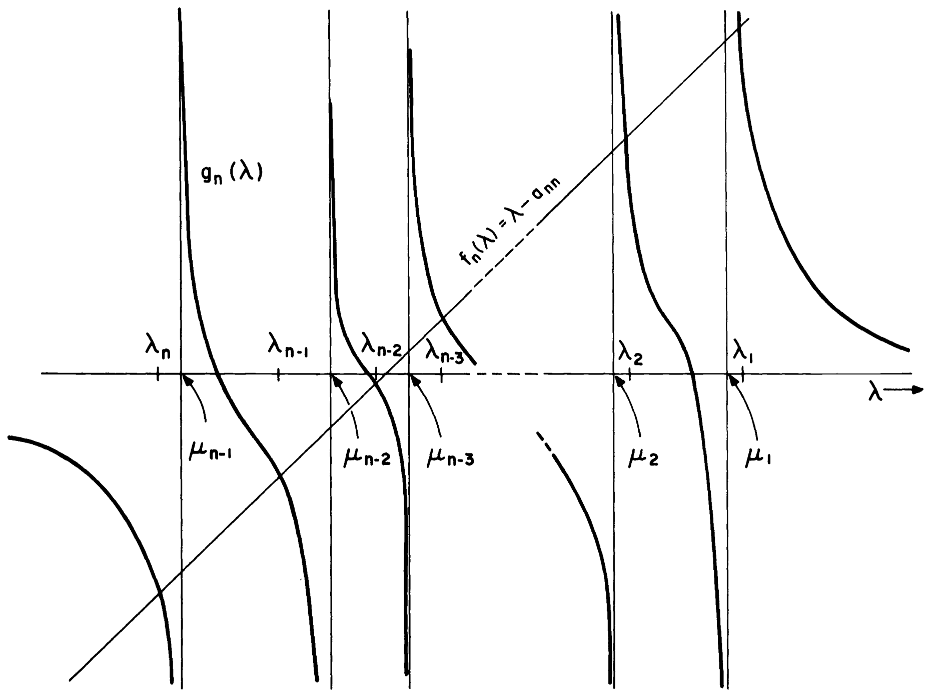

FIG. 1. *Solution of equation* (7).

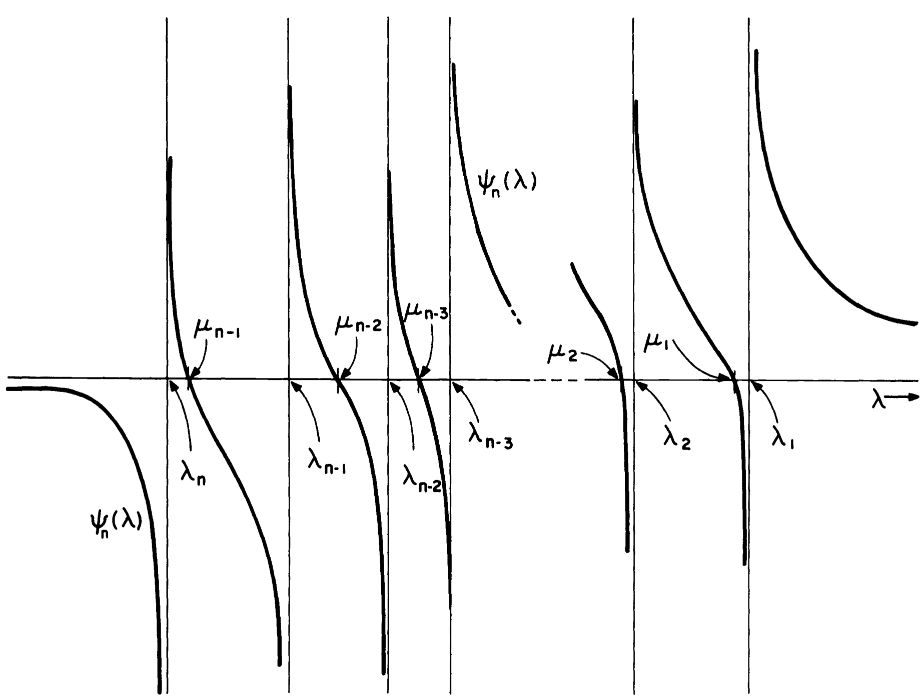

FIG. 2. *Solution of equation* (8).

$k+1$. If $\mathbf{a}$ is orthogonal to all the eigenvectors of B belonging to μ_j and if μ_j is not a root of (7), then μ_j is an eigenvalue of multiplicity k of A. Finally if $\mathbf{a}$ is not orthogonal to all the eigenvectors of B belonging to μ_j, then μ_j is an eigenvalue of multiplicity $k-1$ of A. This classifies the possible ways in which eigenvalues forming the pattern (11) can coalesce.

Similarly, Fig. 2 shows a plot of $\psi_n(\lambda)$ as given in (8) for the case in which none of the quantities $|\alpha_j|, j=1,2,\cdots,n$, is zero and all of the eigenvalues $\lambda_j, j=1,\cdots,n$, of A are of multiplicity one. It is clear again that the separation property (11) then holds. In passing from A to B the multiplicity of an eigenvalue can change by at most one. If λ_j is an eigenvalue of A of multiplicity $k\geqq 1$ and also a root of (8), it is an eigenvalue of B of multiplicity $k+1$. If it does not appear as a pole or root of (8), λ_j is an eigenvalue of B of multiplicity k. If λ_j is a pole of ψ_n it is an eigenvalue of B of multiplicity $k-1$.

Consideration of equation (7) and of Fig. 1 leads readily to a recursively defined upper bound U_n for λ_1 and a recursively defined lower bound L_n for λ_n. Let

(12) $$U_1\equiv L_1\equiv a_{11},$$

(13) $$U_j\equiv\tfrac{1}{2}[a_{jj}+U_{j-1}+\sqrt{(a_{jj}-U_{j-1})^2+4b_j}],$$

(14) $$L_j\equiv\tfrac{1}{2}[a_{jj}+L_{j-1}-\sqrt{(a_{jj}-L_{j-1})^2+4b_j}],$$

(15) $$b_j\equiv\sum_{i=1}^{j-1}|a_{ij}|^2, \qquad j=2,3,\cdots,n.$$

Then it is true that

(16) $$\lambda_1\leqq U_n, \qquad \lambda_n\geqq L_n.$$

These bounds will be established in § 3 below. For diagonal A the bounds are exact. Some numerical examples of the bounds are given in Table 1. Note that the values of U_n and L_n depend on the order in which the rows and columns of A are listed.

TABLE 1

$\lambda_1\leqq U_n, \lambda_n\geqq L_n$. *The bounds* $\hat{U}_n$ *and* $\hat{L}_n$ *are obtained from the matrix with row and column order reversed.* $\|A\|\equiv[\sum a_{ij}^2]^{1/2}$.

Matrix	Eigenvalues	U_n, L_n	$\hat{U}_n$, $\hat{L}_n$
$A=\begin{vmatrix} 1 & .2 & -1 & 0 \\ .2 & 0 & .1 & -.1 \\ -1 & .1 & -.1 & .5 \\ 0 & -.1 & .5 & -2 \end{vmatrix}$	$\lambda_1=1.619$	$U_4=1.6946$	$\hat{U}_4=1.6809$
	$\lambda_2=.0393$		
	$\lambda_3=-.6047$		
	$\lambda_4=-2.1541$	$L_4=-2.2259$	$\hat{L}_4=-2.4351$
	$\|A\|=2.7622$		
$A=\begin{vmatrix} 1 & 0 & 2 & 3 & -1 \\ 0 & 0 & .5 & 1 & -1 \\ 2 & .5 & 0 & -2 & 1 \\ 3 & 1 & -2 & -1 & 0 \\ -1 & -1 & .1 & 0 & -2 \end{vmatrix}$	$\lambda_1=3.5516$	$U_5=5.2448$	$\hat{U}_5=5.5503$
	$\lambda_2=1.5431$		
	$\lambda_3=.2618$		
	$\lambda_4=-2.1955$		
	$\lambda_5=-5.1610$	$L_5=-5.8341$	$\hat{L}_5=-5.8655$
	$\|A\|=6.8206$		
$A=\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 6 & 7 \\ 3 & 6 & 8 & 9 \\ 4 & 7 & 9 & 10 \end{vmatrix}$	$\lambda_1=24.0625$	$U_4=24.0794$	$\hat{U}_4=24.0806$
	$\lambda_2=.5580$		
	$\lambda_3=.1849$		
	$\lambda_4=-.8055$	$L_4=-10.7254$	$\hat{L}_4=-9.7781$
	$\|A\|=24.0832$		

2. Derivation of (7), (8) and (9) and discussion of eigenvalue multiplicity. Suppose that λ is not an eigenvalue of A. Then the system of equations

$$(17)\qquad \begin{aligned} (a_{11}-\lambda)w_1+a_{12}w_2+\cdots+a_{1n}w_n&=0\\ a_{21}w_1+(a_{22}-\lambda)w_2+\cdots+a_{2n}w_n&=0\\ \vdots\qquad\qquad\vdots\qquad\qquad&\vdots\\ a_{n1}w_1+a_{n2}w_2+\cdots+(a_{nn}-\lambda)w_n&=1 \end{aligned}$$

has a unique solution. For w_n, we find by Cramer's rule,

$$(18)\qquad w_n=\frac{\det(B-\lambda I)}{\det(A-\lambda I)}.$$

But (17) can also be written $(A-\lambda I)\mathbf{w}=\mathbf{z}$ where $\mathbf{z}=(0,0,\cdots,1)^T$ so that $\mathbf{w}=(A-\lambda I)^{-1}\mathbf{z}$. It follows that another expression for w_n is

$$(19)\qquad w_n=((A-\lambda I)^{-1}\mathbf{z},\mathbf{z}).$$

Now the eigenvectors $\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_n$ of A are also eigenvectors of $(A-\lambda I)^{-1}$, and corresponding to $\mathbf{x}_j$ is the eigenvalue $(\lambda_j-\lambda)^{-1}$. By expressing $\mathbf{z}$ in terms of the $\mathbf{x}_j$, one finds for the scalar product in (19)

$$(20)\qquad w_n=\sum_1^n\frac{|\alpha_j|^2}{\lambda_j-\lambda}=-\psi_n(\lambda).$$

Consider now the first $n-1$ equations of (17) which can be written

$$(21)\qquad (B-\lambda I)\hat{\mathbf{w}}=-w_n\mathbf{a}$$

where $\mathbf{a}$ is defined by (5) and $\hat{\mathbf{w}}=(w_1,w_2,\cdots,w_{n-1})^T$. The last of equations (17) in this notation is

$$(22)\qquad (\hat{\mathbf{w}},\mathbf{a})+(a_{nn}-\lambda)w_n=1.$$

If now λ is not an eigenvalue of B, (21) can be solved to yield $\hat{\mathbf{w}}=-w_n(B-\lambda I)^{-1}\mathbf{a}$. Inserting this expression for $\hat{\mathbf{w}}$ into (22), we find

$$(23)\qquad \lambda-a_{nn}+((B-\lambda I)^{-1}\mathbf{a},\mathbf{a})=-\frac{1}{w_n}.$$

Since $\mathbf{y}_j$ is an eigenvector of $(B-\lambda I)^{-1}$ corresponding to the eigenvalue $(\mu_j-\lambda)^{-1}$, by expressing $\mathbf{a}$ in terms of the $\mathbf{y}_j$ one can write the scalar product here in the form $\sum(\mu_j-\lambda)^{-1}|\beta_j|^2$, with the β's given by (6). Equations (18), (20) and (23) give (7), (8) and (9).

We now have shown that

$$(24)\qquad \phi_n(\lambda)=\lambda-a_{nn}-\sum_{j=1}^{n-1}\frac{|\beta_j|^2}{\lambda-\mu_j}=\frac{\prod_1^n(\lambda-\lambda_j)}{\prod_1^{n-1}(\lambda-\mu_j)},$$

$$(25)\qquad \psi_n(\lambda)=\sum_1^n\frac{|\alpha_j|^2}{\lambda-\lambda_j}=\frac{\prod_1^{n-1}(\lambda-\mu_j)}{\prod_1^n(\lambda-\lambda_j)}.$$

Singularities of $\phi_n(\lambda)$ and $\psi_n(\lambda)$, if they exist, are here exhibited explicitly as poles of order at most one. Suppose now that λ_j is an eigenvalue of B of multiplicity k and also an eigenvalue of A of multiplicity l. If $k>l+1$, then $\phi_n(\lambda)$ would have a pole of order $\geqq 2$ at $\lambda=\lambda_j$ as is seen from the rightmost member of (24); if $l>k+1$, then $\psi_n(\lambda)$ would have a pole of order $\geqq 2$ as shown by the right member of (25). Since this is impossible, $|k-l|\leqq 1$ and the multiplicity of an eigenvalue can change by at most unity in passing from B to A. The assertions in § 1 following (11) all derive readily from this fact and from (24) and (25).

3. Derivation of inequality (16). Returning to (7), we see that, because of the convention (2),

$$g_n(\lambda)\equiv \sum_{j=1}^{n-1} \frac{|\beta_j|^2}{\lambda-\mu_j} \leqq \sum_{1}^{n-1} \frac{|\beta_j|^2}{\lambda-\mu_1}, \qquad \lambda>\mu_1,$$

$$g_n(\lambda)= \sum_{j=1}^{n-1} \frac{|\beta_j|^2}{\lambda-\mu_j} \geqq \sum_{1}^{n-1} \frac{|\beta_j|^2}{\lambda-\mu_{n-1}}, \qquad \lambda<\mu_{n-1}.$$

Now from (6) it is seen that the quantities β_j and a_{jn}, $j=1,2,\cdots,n-1$, are connected by an $(n-1)\times(n-1)$ unitary transformation, so that

$$\sum_{j=1}^{n-1} |\beta_j|^2 = \sum_{j=1}^{n-1} |a_{jn}|^2.$$

Thus we find

$$g_n(\lambda)\leqq \frac{b_n}{\lambda-\mu_1}, \qquad \lambda>\mu_1,$$

$$g_n(\lambda)\geqq \frac{b_n}{\lambda-\mu_{n-1}}, \qquad \lambda<\mu_{n-1},$$

with b_n defined in (15). Now from (7) and Fig. 1 it follows that $\lambda_1\leqq U_n$ and $\lambda_n\geqq L_n$ where U_n is the largest root of $(\lambda-a_{nn})(\lambda-\mu_1)=b_n$ and L_n is the smallest root of $(\lambda-a_{nn})(\lambda-\mu_{n-1})=b_n$. Thus

$$\lambda_1\leqq U_n=\tfrac{1}{2}[(a_{nn}+\mu_1)+\sqrt{(a_{nn}-\mu_1)^2+4b_n}], \tag{26}$$

$$\lambda_n\geqq L_n=\tfrac{1}{2}[(a_{nn}+\mu_{n-1})-\sqrt{(a_{nn}-\mu_{n-1})^2+4b_n}]. \tag{27}$$

But the right side of (26) is increasing in the quantity μ_1, and the right side of (27) decreases as μ_{n-1} decreases. We apply the equations successively for $n=2,3,\cdots$ to obtain (12)–(16).

Some numerical examples of the use of (12)–(16) are given in Table 1. In general our experience has been that the bounds are tightest when the eigenvalues of A are nearly equally spaced or when A is nearly diagonal. (The bounds are exact for diagonal A.) The first two examples on Table 1 have rather evenly spaced eigenvalues and the bounds are not far from the true eigenvalues. The third example has one large eigenvalue and three small ones: the upper bound U_4 for λ_1 is quite close to the true value, but L_4 is an order of magnitude away from λ_4. The bounds U_n and L_n depend on the order in which the rows and columns of a matrix are listed. On Table 1, the quantities $\hat{U}_n$ and $\hat{L}_n$ are values of U_n and L_n when the rows and columns of the matrix are taken in the opposite order from that listed on the table. For comparison, we also give the quantity $\|A\|\equiv[\sum_1^n |a_{ij}|^2]^{1/2}$, which is a commonly used upper bound for λ_1.

4. Comment. Many of the results given here appear in the literature. See [1, pp. 94–98], for example, for a discussion of the relationship between the quantities λ_j and μ_k. The function $\phi_n(\lambda)$ is used in that discussion. Our inequalities (12)–(16) appear to be new, however, as do also (8) and the curious identity (9). We have thought the approach taken here to be sufficiently novel to warrant repetition of some known material.

Acknowledgment. The authors wish to thank the referee for several helpful comments.

REFERENCES

[1] J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.
[2] R. BELLMAN, *Introduction to Matrix Analysis*, McGraw-Hill, New York, 1970.

SIAM J. MATH. ANAL.
Vol. 9, No. 2, April 1978
Copyright © 1978 Society for Industrial and Applied Mathematics

ON THE BEHAVIOR OF SOLUTIONS OF A SINGULARLY PERTURBED BOUNDARY VALUE PROBLEM WITH A TURNING POINT*

KOICHI NIIJIMA†

Abstract. We find sufficient conditions for the boundary value problem

$$\varepsilon y''(x)+f(x;\varepsilon)y'(x)+g(x;\varepsilon)y(x)=h(x;\varepsilon), \qquad -a\leqq x\leqq b, \tag{1}$$

$$y(-a)=A(\varepsilon), \qquad y(b)=B(\varepsilon) \tag{2}$$

to have a unique solution by applying the Newton–Kantorovič theorem to Riccati equations associated with (1). Moreover, we obtain the behavior of solutions of (1) and (2) by constructing them explicitly.

1. Introduction. In this paper we consider the boundary value problem

$$\varepsilon y''(x)+f(x;\varepsilon)y'(x)+g(x;\varepsilon)y(x)=h(x;\varepsilon), \qquad -a\leqq x\leqq b, \tag{1.1a}$$

$$y(-a)=A(\varepsilon), \qquad y(b)=B(\varepsilon) \tag{1.1b}$$

where $a, b>0$, ε is a small positive parameter, and $f(x;0)$ has a single simple zero at $x=0$ and satisfies $f'(0;0)<0$ (hereafter $x=0$ is called the turning point). The equation (1.1a) has been solved by various asymptotic methods ([2], [7], [11]). Using these methods, several authors ([1], [3], [5], [6], [8], [9], [12], [14]) have recently studied asymptotic approximations to solutions of (1.1). However, the question of exact solutions of (1.1) is not yet sufficiently investigated [4]. In this paper, we study the behavior of exact solutions of (1.1) by constructing them explicitly.

According to Ackerberg and O'Malley [1] and O'Malley [8], asymptotic solutions of the homogeneous problem (1.1) (that is, $h(x;\varepsilon)=0$) can exhibit peculiar behavior which is referred to as "resonance", when $l=-g(0;0)/f'(0;0)$ is a nonnegative integer. Such a phenomenon is of particular interest to us and motivated the present study. When l is not a nonnegative integer, the homogeneous problem (1.1) has a unique solution which is almost zero within $(-a, b)$ [4]. For this reason we restrict our attention to the case where l is a nonnegative integer.

In this paper, applying the Newton–Kantorovič theorem to Riccati equations associated with (1.1a), we find sufficient conditions for the problem (1.1) to have a unique solution and construct the solution to study its behavior. We also show that for the homogenous problem these conditions become sufficient conditions for resonance not to occur. It is further shown that our conditions for resonance not to occur are closely related to those obtained by Cook and Eckhaus [3] and by Kreiss and Parter [4].

2. Preliminaries. In this section we present four lemmas which hold the key to the success of our analysis.

LEMMA 1. *Assume that $\phi(x;\varepsilon)$, $\psi(x,\varepsilon)$, and $\omega(x;\varepsilon)$ are continuous in the region $c\leqq x\leqq d$, $0\leqq\varepsilon\leqq\varepsilon_0$, and that $\phi(x;\varepsilon)$ is positive there. Then the solution $v(x)$ of the initial value problem*

$$\varepsilon v''(x)+\phi(x;\varepsilon)v'(x)+\psi(x;\varepsilon)v(x)=\omega(x;\varepsilon), \qquad c\leqq x\leqq d, \tag{2.1a}$$

$$v(c)=C, \qquad v'(c)=D \tag{2.1b}$$

can be written in the form

$$v(x)=p(x)(D\lambda(x)+\frac{1}{\varepsilon}\mu(x)+C). \tag{2.2}$$

* Received by the editors October 31, 1975, and in revised form October 8, 1976.
† Department of Home Life Science, Fukuoka Women's University, Fukuoka City, Fukuoka, Japan.

Here

$$\lambda(x)=\int_c^x \frac{q(t)}{p^2(t)}\,dt, \qquad \mu(x)=\int_c^x \frac{q(t)}{p^2(t)}\int_c^t \omega(s;\varepsilon)\frac{p(s)}{q(s)}\,ds\,dt,$$

$$q(x)=\exp\left(-\int_c^x \frac{\phi(t;\varepsilon)}{\varepsilon}\,dt\right), \quad \textit{and} \quad p(x)=\exp\left(\int_c^x \alpha(t)\,dt\right)$$

where $\alpha(x)$ *denotes a solution of the Riccati equation*

(2.3a)
$$\alpha'(x)+\frac{\phi(x;\varepsilon)}{\varepsilon}\alpha(x)+\alpha^2(x)+\frac{\psi(x;\varepsilon)}{\varepsilon}=0$$

subject to

(2.3b)
$$\alpha(c)=0$$

Proof. It is well known that the problem (2.1) possesses a unique solution. To obtain the expression (2.2), we shall introduce the linear relation

(2.4)
$$v'(x)=\alpha(x)v(x)+\beta(x).$$

Then it follows from (2.1a) and (2.4) that $\alpha(x)$ satisfies (2.3a) and $\beta(x)$ satisfies

(2.5a)
$$\beta'(x)+\left(\frac{\phi(x;\varepsilon)}{\varepsilon}+\alpha(x)\right)\beta(x)-\frac{\omega(x;\varepsilon)}{\varepsilon}=0.$$

Let $\alpha(c)$ and $\beta(c)$ be chosen as

$$\alpha(c)=0$$

and

(2.5b)
$$\beta(c)=D,$$

respectively. Applying Newton's method to (2.3) and choosing

$$\alpha_0(x)=\int_c^x\left(-\frac{\psi(t;\varepsilon)}{\varepsilon}\right)\exp\left(-\int_t^x \frac{\phi(s;\varepsilon)}{\varepsilon}\,ds\right)dt$$

as an initial iterate of a Newton sequence [10, p. 138], the condition of Yarmish [13, p. 662] is satisfied, so that (2.3) has a solution $\alpha(x)$ satisfying the estimate

$$\|\alpha-\alpha_0\|_\infty \leqq K\varepsilon$$

where $\|\cdot\|_\infty$ denotes the maximum norm and K is a constant independent of ε. Therefore, solving (2.4) and (2.5) directly, we obtain (2.2).

Next we consider solutions of the initial value problem (2.1) for the case where $\phi(x;\varepsilon)$ is not necessarily positive.

LEMMA 2. *Assume that*

(i) $\phi_1(x)$ *and* $\psi_1(x)$ *are Lipschitz continuous on* $[c,d]$,

(ii) $\phi_2(x;\varepsilon)$, $\psi_2(x;\varepsilon)$, *and* $\omega(x;\varepsilon)$ *are continuous in the region* $c\leqq x\leqq d$, $0<\varepsilon\leqq\varepsilon_0$, *and bounded there,*

(iii) $\phi_1(d)=\psi_1(d)=0$,

(iv) $\phi_1(x)>0$ *for* $c\leqq x<d$,

and

(v) $\phi_1'(d)<0$.

Then the solution $v(x)$ of the initial value problem

$$\varepsilon v''(x)+(\phi_1(x)-\varepsilon\phi_2(x;\varepsilon))v'(x)+(\psi_1(x)-\varepsilon\psi_2(x;\varepsilon))v(x)=\omega(x;\varepsilon),\quad c\leqq x\leqq d,$$

$$v(c)=C,\qquad v'(c)=D$$

can also be written in the form

$$v(x)=p(x)\Big(D\lambda(x)+\frac{1}{\varepsilon}\mu(x)+C\Big). \tag{2.6}$$

Here $\lambda(x)$ and $\mu(x)$ have the same expressions as in Lemma 1, *but now*

$$q(x)=\exp\Big(-\int_c^x\Big(\frac{\phi_1(t)}{\varepsilon}-\phi_2(t;\varepsilon)\Big)\,dt\Big)\quad\textit{and}\quad p(x)=\exp\Big(\int_c^x\alpha(t)\,dt\Big)$$

where $\alpha(x)$ is a solution of

$$\alpha'(x)+\Big(\frac{\phi_1(x)}{\varepsilon}-\phi_2(x;\varepsilon)\Big)\alpha(x)+\alpha^2(x)+\frac{\psi_1(x)}{\varepsilon}-\psi_2(x;\varepsilon)=0$$

subject to $\alpha(c)=0$.

Proof. Yarmish's condition does not apply, so we apply Newton's method to the integral equation

$$P(\alpha)(x)=\alpha(x)+\int_c^x\Big(\Big(\frac{\phi_1(t)}{\varepsilon}-\phi_2(t;\varepsilon)\Big)\alpha(t)+\alpha^2(t)+\frac{\psi_1(t)}{\varepsilon}-\psi_2(t;\varepsilon)\Big)\,dt=0, \tag{2.7}$$

and immediately check the conditions of the Newton–Kantorovič theorem [10, p. 135]. Now, as an initial iterate of a Newton sequence $\{\alpha_i\}$ [10, p. 138], we choose

$$\alpha_0(x)=\int_c^x\Big(-\frac{\psi_1(t)}{\varepsilon}+\psi_2(t;\varepsilon)\Big)\exp\Big(-\int_t^x\Big(\frac{\phi_1(s)}{\varepsilon}-\phi_2(s;\varepsilon)\Big)\,ds\Big)\,dt.$$

It follows from (i), (iii), (iv), and (v) that there exist positive constants F_{11}, F_{12} and G_1 satisfying

$$F_{11}(d-x)\leqq\phi_1(x)\leqq F_{12}(d-x)$$

and

$$|\psi_1(x)|\leqq G_1(d-x).$$

Therefore we get

$$\|P'(\alpha_0)^{-1}\|\leqq K_2$$

and

$$\|\alpha_1-\alpha_0\|_\infty\leqq K_3\sqrt{\varepsilon}.$$

Moreover an easy calculation gives

$$\|P''(\tau)\|\leqq 2(d-c)$$

for any $\tau\in C[c,d]$. Therefore (2.7) has a solution $\alpha(x)$ satisfying the estimate

$$\|\alpha-\alpha_0\|_\infty\leqq K\sqrt{\varepsilon} \tag{2.8}$$

where K is a constant independent of ε.

The following two lemmas will be used to find sufficient conditions for the problem (1.1) to have a unique solution.

LEMMA 3. *Let $\alpha(x)$ be a solution of* (2.7).

(i) *If $\phi_1(x)$ and $\psi_1(x)$ belong to $C^2[c, d]$, then*

$$\alpha(d) = -\frac{\psi_1'(d)}{\phi_1'(d)} + O(\sqrt{\varepsilon}). \tag{2.9}$$

(ii) *If $\phi_1(x)$ and $\psi_1(x)$ belong to $C^3[c, d]$, and $\phi_2(x;\varepsilon)$ and $\psi_2(x;\varepsilon)$ are Lipschitz continuous for $c \leqq x \leqq d$, $0 < \varepsilon \leqq \varepsilon_0$, and bounded there, then*

$$\alpha(d) = -\frac{\psi_1'(d)}{\phi_1'(d)} + \left\{ -\left(\frac{\psi_1'(d)}{\phi_1'(d)}\right)^2 + \frac{\phi_1'(d)\psi_1''(d) - \phi_1''(d)\psi_1'(d)}{2[\phi_1'(d)]^2} \right.$$
$$\left. -\frac{\psi_1'(d)}{\phi_1'(d)} \lim_{\varepsilon\to 0} \phi_2(d;\varepsilon) + \lim_{\varepsilon\to 0} \psi_2(d;\varepsilon) \right\}$$
$$\cdot \int_c^d \exp\left(-\int_t^d \left(\frac{\phi_1(s)}{\varepsilon} - \phi_2(s;\varepsilon) + 2\alpha_0(s)\right) ds\right) dt + O(\varepsilon).$$

Proof. Integrating $\alpha_0(x)$ by parts, we get

$$\alpha_0(d) = -\frac{\psi_1'(d)}{\phi_1'(d)} + O(\sqrt{\varepsilon}). \tag{2.10}$$

By combining (2.8) with (2.10), the inequality (2.9) is obtained. The assertion (ii) follows from a property of Newton's sequences,

$$\|\alpha - \alpha_1\|_\infty \leqq C_1 \varepsilon,$$

and partial integration of $\alpha_1(x)$.

LEMMA 4. *Let $\alpha(x)$ be a solution of* (2.3), *where $\phi(x;\varepsilon)$ and $\psi(x;\varepsilon)$ are $l+1$ times continuously differentiable with respect to x in the region $c \leqq x \leqq d$, $0 \leqq \varepsilon \leqq \varepsilon_0$, and $\phi(x;\varepsilon)$ is positive there. Then $p(x) = \exp(\int_c^x \alpha(t)\,dt)$ satisfies the estimates*

$$|p^{(j)}(x) - \tilde{p}^{(j)}(x)| \leqq K\left(\varepsilon^{-\max(j-1,0)} \exp\left(-\frac{m}{\varepsilon}(x-c)\right) + \varepsilon\right), \tag{2.11}$$
$$j = 0, 1, 2, \cdots, l+1,$$

where

$$\tilde{p}(x) = \exp\left(-\int_c^x \frac{\psi(t;0)}{\phi(t;0)} dt\right)$$

and K is a constant independent of ε.

Proof. Let $\tilde{\alpha}(x) = -\psi(x;0)/\phi(x;0)$ and $\gamma(x) = \alpha(x) - \tilde{\alpha}(x)$. Then we show by induction that the estimates

$$|\gamma^{(j)}(x)| \leqq K_j\left(\varepsilon^{-j} \exp\left(-\frac{m}{\varepsilon}(x-c)\right) + \varepsilon\right), \qquad j = 0, 1, \cdots, l, \tag{2.12}$$

and

$$|\alpha^{(j+1)}(x)| \leqq C_{j+1}\left(\varepsilon^{-(j+1)} \exp\left(-\frac{m}{\varepsilon}(x-c)\right) + 1\right), \qquad j = 0, 1, \cdots, l, \tag{2.13}$$

hold for $c \leqq x \leqq d$. Here K_j and C_{j+1} are constants independent of ε and

$$m = \min_{\substack{c \leqq x \leqq d \\ 0 \leqq \varepsilon \leqq \varepsilon_0}} \phi(x;\varepsilon).$$

First, it follows from (2.3a) and the definition of $\tilde{\alpha}(x)$ that $\gamma(x)$ satisfies

$$(2.14) \qquad \begin{aligned} &\gamma'(x)+\frac{\phi(x;\varepsilon)}{\varepsilon}\gamma(x)+\frac{\phi(x;\varepsilon)\tilde{\alpha}(x)+\psi(x;\varepsilon)}{\varepsilon}+\tilde{\alpha}'(x)+\alpha^2(x)=0,\\ &\gamma(c)=\frac{\psi(c;0)}{\phi(c;0)}. \end{aligned}$$

Solving this, we obtain the estimate

$$|\gamma(x)| \leqq K_0\left(\exp\left(-\frac{m}{\varepsilon}(x-c)\right)+\varepsilon\right)$$

for $c \leqq x \leqq d$. Furthermore, by using (2.14) and the equation $\alpha'(x)=\tilde{\alpha}'(x)+\gamma'(x)$, we get

$$|\alpha'(x)| \leqq C_1\left(\varepsilon^{-1}\exp\left(-\frac{m}{\varepsilon}(x-c)\right)+1\right).$$

Next we assume that (2.12) and (2.13) hold for $j=0, 1, 2, \cdots, k-1$. Differentiating (2.14) k times with respect to x, we obtain

$$(2.15) \qquad \begin{aligned} \gamma^{(k+1)}(x)+\frac{\phi(x;\varepsilon)}{\varepsilon}\gamma^{(k)}(x) &= -\frac{1}{\varepsilon}\sum_{j=0}^{k-1}\binom{k}{j}\phi^{(k-j)}(x;\varepsilon)\gamma^{(j)}(x)\\ &\quad -\delta^{(k)}(x)-\sum_{j=0}^{k}\binom{k}{j}\alpha^{(k-j)}(x)\alpha^{(j)}(x) \end{aligned}$$

where $\delta(x)=(\phi(x;\varepsilon)\tilde{\alpha}(x)+\psi(x;\varepsilon))/\varepsilon+\tilde{\alpha}'(x)$. By the inductive assumptions, we see that the right hand side $r(x)$ of (2.15) satisfies the estimate

$$|r(x)| \leqq R\left(\varepsilon^{-k}\exp\left(-\frac{m}{\varepsilon}(x-c)\right)+1\right)$$

for $c \leqq x \leqq d$, where R is a constant independent of ε. On the other hand, differentiating (2.14) $k-1$ times with respect to x and again using the inductive assumptions, we get

$$|\gamma^{(k)}(c)| \leqq K_k'\varepsilon^{-k}$$

where K_k' is a constant independent of ε. Therefore $\gamma^{(k)}(x)$ satisfies

$$|\gamma^{(k)}(x)| \leqq K_k\left(\varepsilon^{-k}\exp\left(-\frac{m}{\varepsilon}(x-c)\right)+\varepsilon\right)$$

for $c \leqq x \leqq d$. This estimate together with (2.15) implies that

$$|\alpha^{(k+1)}(x)| \leqq C_{k+1}\left(\varepsilon^{-(k+1)}\exp\left(-\frac{m}{\varepsilon}(x-c)\right)+1\right).$$

Thus (2.12) and (2.13) follow by induction. The inequality (2.11) can now be proved by using (2.12), (2.13), and the equation $p'(x)-\tilde{p}'(x)=\alpha(x)(p(x)-\tilde{p}(x))+\gamma(x)\tilde{p}(x)$.

3. Main results. In this section we find sufficient conditions for the problem (1.1) to have a unique solution, and construct the solution explicitly by using the four lemmas obtained in the previous section. We shall assume that $f(x;\varepsilon)$ and $g(x;\varepsilon)$ are Lipschitz continuous with respect to x and ε in the region $-a \leqq x \leqq b$, $0 \leqq \varepsilon \leqq \varepsilon_0$, and $h(x;\varepsilon)$ is continuous there. Assume that $A(\varepsilon)$ and $B(\varepsilon)$ are continuous on $[0, \varepsilon_0]$. First we consider the following case.

(a) The case when $l=0$. Instead of the problem (1.1), we consider two initial value problems

$$\varepsilon y''(x)+f(x;\varepsilon)y'(x)+g(x;\varepsilon)y(x)=h(x;\varepsilon), \qquad -a\leqq x\leqq 0, \tag{3.1a}$$

$$y(-a)=A(\varepsilon), \qquad y'(-a)=X \tag{3.1b}$$

and

$$\varepsilon y''(x)+f(x;\varepsilon)y'(x)+g(x;\varepsilon)y(x)=h(x;\varepsilon), \qquad 0\leqq x\leqq b, \tag{3.2a}$$

$$y(b)=B(\varepsilon), \qquad y'(b)=Y \tag{3.2b}$$

containing two unknown numbers X and Y. For the problem (3.1), since the assumptions of Lemma 2 are satisfied if we set $\phi_1(x)=f(x;0)$, $\psi_1(x)=g(x;0)$, $\phi_2(x;\varepsilon)=(f(x;0)-f(x;\varepsilon))/\varepsilon$, $\psi_2(x;\varepsilon)=(g(x;0)-g(x;\varepsilon))/\varepsilon$, and $\omega(x;\varepsilon)=h(x;\varepsilon)$, the solution of (3.1) can be written in the form (2.6) on $[-a, 0]$, where $C=A(\varepsilon)$ and $D=X$. We denote this by $y_-(x)$, the functions $\alpha(x)$, $p(x)$, $q(x)$, $\lambda(x)$, *and* $\mu(x)$ in $y_-(x)$ by $\alpha_-(x)$, $p_-(x)$, $q_-(x)$, $\lambda_-(x)$, and $\mu_-(x)$, respectively. For the problem (3.2), if we apply the change of variable $z=b-x$ and set $\phi_1(x)=-f(b-x;0)$, $\psi_1(x)=g(b-x;0)$, $\phi_2(x;\varepsilon)=(f(b-x;\varepsilon)-f(b-x;0))/\varepsilon$, $\psi_2(x;\varepsilon)=(g(b-x;0)-g(b-x;\varepsilon))/\varepsilon$, and $\omega(x;\varepsilon)=h(b-x;\varepsilon)$, then the assumptions of Lemma 2 are satisfied; therefore $w(x)=y(b-x)$ can be written in the form (2.6) on $[0, b]$, where $C=B(\varepsilon)$ and $D=-Y$. Here we denote $\alpha(x)$, $p(x)$, $q(x)$, $\lambda(x)$, and $\mu(x)$ in $w(x)$ by $\alpha_+(x)$, $p_+(x)$, $q_+(x)$, $\lambda_+(b-x)$, and $\mu_+(b-x)$, respectively, and set $y_+(x)=w(b-x)$, which becomes a solution of (3.2). If we can determine X and Y so as to satisfy

$$y_-(0)=y_+(0) \tag{3.3a}$$

and

$$y'_-(0)=y'_+(0), \tag{3.3b}$$

then

$$y(x)=\begin{cases} y_-(x), & -a\leqq x\leqq 0,\\ y_+(x), & 0\leqq x\leqq b,\end{cases}$$

becomes a solution of (1.1).

Then we have the following theorem.

THEOREM 1. *Let us assume that*

(i) $f(x;0)$ *and* $g(x;0)$ *are two times continuously differentiable on* $[-a, 0]$ *and* $[0, b]$, *and* $g'_-(0;0)\neq g'_+(0;0)$ *holds, where* $g'_-(0;0)$ *and* $g'_+(0;0)$ *denote the left and the right derivatives of* $g(x;0)$ *at* $x=0$, *respectively,*

or

(ii) $f(x;0)$ *and* $g(x;0)$ *belong to* $C^2[-a, b]$ *and are three times continuously differentiable on* $[-a, 0]$ *and* $[0, b]$, *and* $(\partial^j f/\partial\varepsilon^j)(x;\varepsilon)$ *and* $(\partial^j g/\partial\varepsilon^j)(x;\varepsilon)$, $j=0, 1, 2$, *are continuous in the region* $-a\leqq x\leqq b$, $0\leqq\varepsilon\leqq\varepsilon_0$, *and moreover* $(\partial^3 f/\partial\varepsilon^3)(x;\varepsilon)$ *and* $(\partial^3 g/\partial\varepsilon^3)(x;\varepsilon)$ *are continuous in the subregions* $-a\leqq x\leqq 0$,

$0 \leqq \varepsilon \leqq \varepsilon_0$ *and* $0 \leqq x \leqq b, 0 \leqq \varepsilon \leqq \varepsilon_0$, *and in addition* $f(x; \varepsilon)$ *and* $g(x; \varepsilon)$ *satisfy*

$$I = -\left(\frac{g'(0; 0)}{f'(0; 0)}\right)^2 + \frac{f'(0; 0)g''(0; 0) - f''(0; 0)g'(0; 0)}{2[f'(0; 0)]^2} + \frac{\partial f}{\partial \varepsilon}(0; 0)\frac{g'(0; 0)}{f'(0; 0)} - \frac{\partial g}{\partial \varepsilon}(0; 0) \neq 0.$$

Then, for sufficiently small $\varepsilon > 0$, *the problem* (1.1) *has a unique solution and the solution can be written as*

$$y(x) = \begin{cases} A(\varepsilon)p_-(x)\left(1 - \dfrac{\lambda_-(x)}{\lambda_-(0)}\right) \\ \quad + \dfrac{p_-(x)}{\varepsilon}\left\{\mu_-(x) - \left(\mu_-(0) + \dfrac{\mu'_-(0)p_-(0) - \mu'_+(0)p_+(b)}{(\alpha_-(0) + \alpha_+(b))p_-(0)}\right)\dfrac{\lambda_-(x)}{\lambda_-(0)}\right\} + T_1, \\ \hfill -a \leqq x \leqq 0, \\ B(\varepsilon)p_+(b - x)\left(1 - \dfrac{\lambda_+(x)}{\lambda_+(0)}\right) \\ \quad + \dfrac{p_+(b - x)}{\varepsilon}\left\{\mu_+(x) - \left(\mu_+(0) + \dfrac{\mu'_-(0)p_-(0) - \mu'_+(0)p_+(b)}{(\alpha_-(0) + \alpha_+(b))p_+(b)}\right)\dfrac{\lambda_+(x)}{\lambda_+(0)}\right\} + T_2, \\ \hfill 0 \leqq x \leqq b. \end{cases}$$

Here, for $j = 1, 2$,

$$T_j = \begin{cases} O\left(\varepsilon^{-1}\exp\left(-\dfrac{\kappa}{\varepsilon}\right)\right), & \text{if (i) is satisfied,} \\ O\left(\varepsilon^{-3/2}\exp\left(-\dfrac{\kappa}{\varepsilon}\right)\right), & \text{if (ii) is satisfied,} \end{cases}$$

where $\kappa = \min\left(\int_{-a}^{0} f(t; \varepsilon)\,dt, -\int_0^b f(t; \varepsilon)\,dt\right)$. *Moreover,* $y'(-a)$ *and* $y'(b)$ *are*

$$y'(-a) = -\frac{A(\varepsilon) + \dfrac{1}{\varepsilon}\left\{\mu_-(0) + \dfrac{\mu'_-(0)p_-(0) - \mu'_+(0)p_+(b)}{(\alpha_-(0) + \alpha_+(b))p_-(0)}\right\}}{\lambda_-(0)} + R_1$$

and

$$y'(b) = \frac{B(\varepsilon) + \dfrac{1}{\varepsilon}\left\{\mu_+(0) + \dfrac{\mu'_-(0)p_-(0) - \mu'_+(0)p_+(b)}{(\alpha_-(0) + \alpha_+(b))p_+(b)}\right\}}{\lambda_+(0)} + R_2,$$

respectively, where, for $j = 1, 2$,

$$R_j = \begin{cases} O\left(\varepsilon^{-3/2}\exp\left(-\dfrac{\kappa}{\varepsilon}\right)\right), & \text{if (i) is satisfied,} \\ O\left(\varepsilon^{-2}\exp\left(-\dfrac{\kappa}{\varepsilon}\right)\right), & \text{if (ii) is satisfied.} \end{cases}$$

Proof. The determinant of the coefficient matrix M of (3.3) is given by

$$\det M = (\alpha_-(0)+\alpha_+(b))p_-(0)p_+(b)\lambda_-(0)\lambda_+(0)+q_-(0)\frac{p_+(b)}{p_-(0)}\lambda_+(0) \tag{3.4}$$
$$+q_+(b)\frac{p_-(0)}{p_+(b)}\lambda_-(0).$$

From this, we see that for sufficiently small $\varepsilon > 0$, $\det M$ is governed by the first term of the right hand side of (3.4) which can be evaluated by using Lemma 3. Indeed, applying Lemma 3 to $\alpha_-(x)$ and $\alpha_+(x)$, we obtain

$$\alpha_-(0)+\alpha_+(b) = \frac{g'_+(0;0)-g'_-(0;0)}{f'(0;0)}+O(\sqrt{\varepsilon}) = O(1)$$

if (i) is satisfied, and

$$\alpha_-(0)+\alpha_+(b) = I\left\{\int_{-a}^{0} \exp\left(-\int_t^0 \left(\frac{f(s;\varepsilon)}{\varepsilon}+2\alpha_{-,0}(s)\right)ds\right)dt\right.$$
$$\left.+\int_0^b \exp\left(\int_t^b \left(\frac{f(b-s;\varepsilon)}{\varepsilon}-2\alpha_{+,0}(s)\right)ds\right)dt\right\}+O(\varepsilon) = O(\sqrt{\varepsilon})$$

if (ii) is satisfied. Thus, for sufficiently small $\varepsilon > 0$, $\det M \neq 0$. Therefore the problem (1.1) has a solution. The uniqueness of the solution follows from the Fredholm alternative theorem. The quantities $X = y'(-a)$ and $Y = y'(b)$ are obtained by solving (3.3). Substituting these into $y_-(x)$ and $y_+(x)$, we get the solution $y(x)$. This completes the proof.

From these results, one finds that when (1.1a) is homogeneous the solution is almost zero within $(-a, b)$ and its slope at both endpoints is $O(\varepsilon^{-1/2})$. When (1.1a) is inhomogeneous, the solution is $O(\varepsilon^{-1/2})$ within $(-a, b)$ if (i) is satisfied, and $O(\varepsilon^{-1})$ within $(-a, b)$ if (ii) is satisfied.

Remark 1. Suppose that $f(x;\varepsilon) = -xa(x)$ where $a(x) \geqq a_0 > 0$ and $a(x) \in C^2[-a, b]$, and that $g(x;\varepsilon) = x^2 b(x)$ where $b(x) \geqq b_0 > 0$ and $b(x) \in C^2[-a, b]$. Then we have

$$I = -b(0)/a(0) \neq 0.$$

Therefore Theorem 3.1 of Kreiss and Parter [4, p. 242] is a particular case of our Theorem 1.

In Theorem 3.2 [4, p. 243], if (e) is not assumed, the condition that $g^{(k)}(0) \neq 0$ for some positive integer k implies that $g''(0) \neq 0$. This condition is equivalent to $I \neq 0$, since $I = -g''(0)/2$.

The latter part of Theorem 3.3 [4, p. 244] will be shown in § 4 of this paper.

Remark 2. The quantity I is related to the quantity σ with $l = 0$ defined by Cook and Eckhaus [3, p. 135] as follows:

$$I = -f'(0;0)\sigma.$$

Thus, for the case of $l = 0$, their criterion for resonance ($\sigma = 0$) is obtained by our method.

To get improved conditions for resonance, we must calculate the higher order terms on ε of $\alpha_-(0)+\alpha_+(b)$ under the assumptions that $f(x;\varepsilon)$ and $g(x;\varepsilon)$ are sufficiently smooth.

Next we consider the following case.

(b) The case when l is a positive integer. In this case, we assume that the condition below will hold:

H1. $f(x;\varepsilon)$, $g(x;\varepsilon)$, and $h(x;\varepsilon)$ are l times continuously differentiable with respect to x for $-\Delta \leqq x \leqq \Delta$, $0 \leqq \varepsilon \leqq \varepsilon_0$, where Δ is a positive number independent of ε.

Since $g(0;0) \geqq -f'(0;0) > 0$, there exists a positive number $\tilde{\delta}$ such that $g(x;\varepsilon)$ is positive in the region $-\tilde{\delta} \leqq x \leqq \tilde{\delta}$, $0 \leqq \varepsilon \leqq \varepsilon_0$. Let δ satisfy $0<\delta<\min(\Delta, \tilde{\delta})$. We analyze the problem (1.1) on each of subintervals $[-a, -\delta]$, $[-\delta, \delta]$, and $[\delta, b]$. In the interval $[-a, -\delta]$, applying Lemma 1 to (1.1a) subject to $y(-a)=A(\varepsilon)$ and $y'(-a)=X$, we obtain a solution $y_-(x)$ of the form (2.2). Here we denote $\alpha(x)$, $p(x)$, $q(x)$, $\lambda(x)$, and $\mu(x)$ in $y_-(x)$ by $\alpha_-(x)$, $p_-(x)$, $q_-(x)$, $\lambda_-(x)$, and $\mu_-(x)$, respectively. In the interval $[\delta, b]$, applying the change of variable $z=b-x$ to (1.1a) and thereafter using Lemma 1, we can get a function $w(x)=y(b-x)$ of the form (2.2) on $[0, b-\delta]$. Here we denote $\alpha(x), p(x), q(x), \lambda(x)$, and $\mu(x)$ in $w(x)$ by $\alpha_+(x), p_+(x), q_+(x), \lambda_+(b-x)$, and $\mu_+(b-x)$, respectively, and set $y_+(x)=w(b-x)$, which becomes a solution of (1.1a) subject to $y(b)=B(\varepsilon)$ and $y'(b)=Y$ on $[\delta, b]$. In the remaining interval $[-\delta, \delta]$, we differentiate (1.1a) l times with respect to x, which leads us to the following differential equation for $v_l(x)=y^{(l)}(x)$:

$$\varepsilon v_l''(x)+(f(x;0)-\varepsilon f_l(x;\varepsilon))v_l'(x)+(lf'(x;0)+g(x;0) \\ -k_l(x;\varepsilon)f(x;0)-\varepsilon g_l(x;\varepsilon))v_l(x)=h_l(x;\varepsilon). \tag{3.5}$$

Here $f_l(x;\varepsilon)$, $k_l(x;\varepsilon)$, $g_l(x;\varepsilon)$, and $h_l(x;\varepsilon)$ are determined by the iteration formulas

$$f_{j+1}(x;\varepsilon)=f_j(x;\varepsilon)+r_j(x;\varepsilon),$$
$$k_{j+1}(x;\varepsilon)=k_j(x;\varepsilon)+r_j(x;\varepsilon),$$
$$g_{j+1}(x;\varepsilon)=g_j(x;\varepsilon)+f_j'(x;\varepsilon)-r_j(x;\varepsilon)f_j(x;\varepsilon),$$

and

$$h_{j+1}(x;\varepsilon)=h_j'(x;\varepsilon)-r_j(x;\varepsilon)h_j(x;\varepsilon),$$

having $f_0(x;\varepsilon)=(f(x;0)-f(x;\varepsilon))/\varepsilon$, $k_0(x;\varepsilon)=0$, $g_0(x;\varepsilon)=(g(x;0)-g(x;\varepsilon))/\varepsilon$, and $h_0(x;\varepsilon)=h(x;\varepsilon)$ as initial functions, respectively, where

$$r_j(x;\varepsilon)=\frac{jf''(x;0)+g'(x;0)-k_j'(x;\varepsilon)f(x;0)-k_j(x;\varepsilon)f'(x;0)-\varepsilon g_j'(x;\varepsilon)}{jf'(x;0)+g(x;0)-k_j(x;\varepsilon)f(x;0)-\varepsilon g_j(x;\varepsilon)}.$$

From these formulas, we see that $f_l(x;\varepsilon)$, $k_l(x;\varepsilon)$, $g_l(x;\varepsilon)$, and $h_l(x;\varepsilon)$ are continuous for $-\delta \leqq x \leqq \delta$, $0<\varepsilon \leqq \varepsilon_0$, and bounded there. We further assume

H2. $f'(x;0)$ and $\lim_{\varepsilon\to 0} k_l(x;\varepsilon)$ are Lipschitz continuous for $-\delta \leqq x \leqq \delta$ and $k_l(x;\varepsilon)$ is continuously differentiable with respect to ε in the region $-\delta \leqq x \leqq \delta$, $0<\varepsilon \leqq \varepsilon_0$, and its derivative is bounded there.

Then, if we set

$$
(3.6)\qquad
\begin{aligned}
&\phi_1(x)=f(x;0),\\
&\psi_1(x)=lf'(x;0)+g(x;0)-\lim_{\varepsilon\to 0}k_l(x;\varepsilon)\cdot f(x;0),\\
&\phi_2(x;\varepsilon)=f_l(x;\varepsilon),\\
&\psi_2(x;\varepsilon)=\frac{k_l(x;\varepsilon)-\lim_{\varepsilon\to 0}k_l(x;\varepsilon)}{\varepsilon}f(x;0)+g_l(x;\varepsilon),
\end{aligned}
$$

and

$$\omega(x;\varepsilon)=h_l(x;\varepsilon),$$

it follows from Lemma 2 that the solution of (3.5) subject to $v_l(-\delta)=y_-^{(l)}(-\delta)$ and $v_l'(-\delta)=y_-^{(l+1)}(-\delta)$ can be written as (2.6) on $[-\delta,0]$. We denote this by $v_{l,-}(x)$, the functions $\alpha(x)$, $p(x)$, $q(x)$, $\lambda(x)$, and $\mu(x)$ in $v_{l,-}(x)$ by $\hat{\alpha}_-(x)$, $\hat{p}_-(x)$, $\hat{q}_-(x)$, $\hat{\lambda}_-(x)$, and $\hat{\mu}_-(x)$, respectively. In the interval $[0,\delta]$, applying the change of variable $z=\delta-x$ to (3.5) and thereafter using Lemma 2, we can obtain a function $w(x)=v_l(\delta-x)$ of the form (2.6) on $[0,\delta]$. Here we denote $\alpha(x)$, $p(x)$, $q(x)$, $\lambda(x)$, and $\mu(x)$ in $w(x)$ by $\hat{\alpha}_+(x)$, $\hat{p}_+(x)$, $\hat{q}_+(x)$, $\hat{\lambda}_+(\delta-x)$, and $\hat{\mu}_+(\delta-x)$, respectively, and set $v_{l,+}(x)=w(\delta-x)$, which becomes a solution of (3.5) subject to $v_l(\delta)=y_+^{(l)}(\delta)$ and $v_l'(\delta)=y_+^{(l+1)}(\delta)$ on $[0,\delta]$. We shall determine X and Y to satisfy the matching conditions

$$(3.7a)\qquad v_{l,-}(0)=v_{l,+}(0)$$

and

$$(3.7b)\qquad v_{l,-}'(0)=v_{l,+}'(0).$$

We have the following theorem.

THEOREM 2. *Assume that $f(x;0)$, $g(x;0)$, $f_l(x;\varepsilon)$, $k_l(x;\varepsilon)$, and $g_l(x;\varepsilon)$ satisfy the conditions*

(i) *$f(x;0)$ and $g(x;0)$ belong to $C^4[-\delta,\delta]$ and $C^3[-\delta,\delta]$, respectively, and $\lim_{\varepsilon\to 0}k_l(x;\varepsilon)$ is of class $C^3[-\delta,\delta]$,*

(ii) *$f_l(x;\varepsilon)$, $k_l(x;\varepsilon)$, and $g_l(x;\varepsilon)$ are continuously differentiable with respect to x and ε in the region $-\delta\leqq x\leqq\delta$, $0<\varepsilon\leqq\varepsilon_0$, and their derivatives are bounded there, and $(\partial^2k_l/\partial x\partial\varepsilon)(x;\varepsilon)$ are continuous for $-\delta\leqq x\leqq\delta$, $0<\varepsilon\leqq\varepsilon_0$, and bounded there,*

and

(iii)
$$
\begin{aligned}
J=&\frac{\partial f}{\partial\varepsilon}\frac{g'}{f'}+l\frac{\partial f}{\partial\varepsilon}\frac{f''}{f'}-\frac{6l+1}{2}\frac{f''g'}{f'^2}-\frac{3l^2}{2}\left(\frac{f''}{f'}\right)^2\\
&-\left(\frac{g'}{f'}\right)^2+\frac{2l+1}{2}\frac{g''}{f'}-\frac{\partial g}{\partial\varepsilon}-l\frac{\partial^2 f}{\partial x\partial\varepsilon}+\frac{l^2}{2}\frac{f'''}{f'}\\
&\neq 0,
\end{aligned}
$$

where f', f'', and f''' denote the first, the second, and the third derivatives of $f(x;0)$ at $x=0$, respectively, g' and g'' denote the first and the second derivatives of $g(x;0)$ at $x=0$, respectively, and $\partial f/\partial\varepsilon$, $\partial g/\partial\varepsilon$, and $\partial^2f/\partial x\partial\varepsilon$ mean $(\partial f/\partial\varepsilon)(0;0)$, $(\partial g/\partial\varepsilon)(0;0)$, and $(\partial^2f/\partial x\partial\varepsilon)(0;0)$, respectively.

Then, for sufficiently small $\varepsilon>0$, the problem (1.1) *has a unique solution and the solution can be written as*

$$y(x)=\begin{cases} p_-(x)\Big(X\lambda_-(x)+\frac{1}{\varepsilon}\mu_-(x)+A(\varepsilon)\Big), & -a\leqq x\leqq -\delta,\\ \sum_{j=0}^{l-1}\frac{1}{j!}y_-^{(j)}(-\delta)(x+\delta)^j \\ \quad +\frac{1}{(l-1)!}\int_{-\delta}^{x}(x-t)^{l-1}\hat{p}_-(t)\Big(y_-^{(l+1)}(-\delta)\hat{\lambda}_-(t)+\frac{1}{\varepsilon}\hat{\mu}_-(t)+y_-^{(l)}(-\delta)\Big)\,dt, \\ & -\delta\leqq x\leqq 0,\\ \sum_{j=0}^{l-1}\frac{1}{j!}y_+^{(j)}(\delta)(x-\delta)^j \\ \quad +\frac{1}{(l-1)!}\int_{\delta}^{x}(x-t)^{l-1}\hat{p}_+(\delta-t)\Big(-y_+^{(l+1)}(\delta)\hat{\lambda}_+(t)+\frac{1}{\varepsilon}\hat{\mu}_+(t)+y_+^{(l)}(\delta)\Big)\,dt, \\ & 0\leqq x\leqq \delta,\\ p_+(b-x)\Big(-Y\lambda_+(x)+\frac{1}{\varepsilon}\mu_+(x)+B(\varepsilon)\Big), & \delta\leqq x\leqq b. \end{cases}$$

Here

$$X=-\frac{b_1+\hat{\lambda}_-(0)b_2}{a_1+\hat{\lambda}_-(0)a_2}A(\varepsilon)-\frac{\xi_1+\hat{\lambda}_-(0)\xi_2+\hat{\mu}_-(0)}{\varepsilon(a_1+\hat{\lambda}_-(0)a_2)} -\frac{\hat{\mu}'_-(0)\hat{p}_-(0)-\hat{\mu}'_+(0)\hat{p}_+(\delta)}{\varepsilon(a_1+\hat{\lambda}_-(0)a_2)(\hat{\alpha}_-(0)+\hat{\alpha}_+(\delta))\hat{p}_-(0)}+O\Big(\varepsilon^{-(l+3/2)}\exp\Big(-\frac{\kappa_\delta}{\varepsilon}\Big)\Big)$$

and

$$Y=\frac{d_1-\hat{\lambda}_+(0)d_2}{c_1-\hat{\lambda}_+(0)c_2}B(\varepsilon)+\frac{\eta_1-\hat{\lambda}_+(0)\eta_2+\hat{\mu}_+(0)}{\varepsilon(c_1-\hat{\lambda}_+(0)c_2)} +\frac{\hat{\mu}'_-(0)\hat{p}_-(0)-\hat{\mu}'_+(0)\hat{p}_+(\delta)}{\varepsilon(c_1-\hat{\lambda}_+(0)c_2)(\hat{\alpha}_-(0)+\hat{\alpha}_+(\delta))\hat{p}_+(\delta)}+O\Big(\varepsilon^{-(l+3/2)}\exp\Big(-\frac{\kappa_\delta}{\varepsilon}\Big)\Big),$$

where a_j, b_j, ξ_j, c_j, d_j, and η_j denote $(p_-(x)\lambda_-(x))^{(l-1+j)}|_{x=-\delta}$, $p_-^{(l-1+j)}(-\delta)$, $(p_-(x)\mu_-(x))^{(l-1+j)}|_{x=-\delta}$, $(p_+(b-x)\lambda_+(x))^{(l-1+j)}|_{x=\delta}$, $(-1)^{l-1+j}p_+^{(l-1+j)}(b-\delta)$, and $(p_+(b-x)\mu_+(x))^{(l-1+j)}|_{x=\delta}$, respectively, and $\kappa_\delta=\min\,(\int_{-\delta}^{0}f(t;\varepsilon)\,dt,\ -\int_0^{\delta}f(t;\varepsilon)\,dt)$.

Proof. The determinant of the coefficient matrix N of (3.7) is

$$\det N=-(\hat{\alpha}_-(0)+\hat{\alpha}_+(\delta))\hat{p}_-(0)\hat{p}_+(\delta)(a_1+a_2\hat{\lambda}_-(0))(c_1-c_2\hat{\lambda}_+(0)) +\hat{q}_+(\delta)\frac{\hat{p}_-(0)}{\hat{p}_+(\delta)}(a_1+a_2\hat{\lambda}_-(0))c_2-\hat{q}_-(0)\frac{\hat{p}_+(\delta)}{\hat{p}_-(0)}(c_1-c_2\hat{\lambda}_+(0))a_2.$$

Moreover, applying Lemma 4 to $p_-(x)$ and $p_+(b-x)$ of a_j and c_j, we obtain

$$\det N=(-1)^{l+1}(\hat{\alpha}_-(0)+\hat{\alpha}_+(\delta))\hat{p}_-(0)\hat{p}_+(\delta)(\tilde{p}_-^{(l)}(-\delta)\lambda_-(-\delta)+O(\varepsilon^{3/2})) \cdot(\tilde{p}_+^{(l)}(b-\delta)\lambda_+(\delta)+O(\varepsilon^{3/2}))+O\Big(\varepsilon^2\exp\Big(-\frac{\kappa_\delta}{\varepsilon}\Big)\Big).$$

Note that $\tilde{p}_-^{(l)}(-\delta)=O(1)$ and $\tilde{p}_+^{(l)}(b-\delta)=O(1)$. By Lemma 3,

$$\hat{\alpha}_-(0)+\hat{\alpha}_+(\delta)=I\left\{\int_{-\delta}^{0}\exp\left(-\int_t^0\left(\frac{\phi_1(s)}{\varepsilon}-\phi_2(s;\varepsilon)+2\alpha_{-,0}(s)\right)ds\right)dt\right.$$
$$\left.+\int_0^{\delta}\exp\left(\int_t^{\delta}\left(\frac{\phi_1(\delta-s)}{\varepsilon}+\phi_2(\delta-s;\varepsilon)-2\alpha_{+,0}(s)\right)ds\right)dt\right\},$$

where the quantity I is

$$I=-\left(\frac{\phi_1'(0)}{\psi_1'(0)}\right)^2+\frac{\phi_1'(0)\psi_1''(0)-\phi_1''(0)\psi_1'(0)}{2[\phi_1'(0)]^2}-\frac{\phi_1'(0)}{\psi_1'(0)}\lim_{\varepsilon\to 0}\phi_2(0;\varepsilon)+\lim_{\varepsilon\to 0}\psi_2(0;\varepsilon)$$

and $\phi_1(x)$, $\psi_1(x)$, $\phi_2(x;\varepsilon)$, and $\psi_2(x;\varepsilon)$ denote the functions defined by (3.6), respectively. Using the iteration formulas of $f_j(x;\varepsilon)$, $k_j(x;\varepsilon)$, and $g_j(x;\varepsilon)$, we can further calculate the quantity I. Indeed, a direct calculation gives

$$k_l(0;0)=-\frac{l(l-1)}{2}\frac{f''}{f'}-l\frac{g'}{f'},$$

$$\frac{\partial k_l}{\partial x}(0;0)=-\frac{l(l-1)(l+4)}{12}\frac{f'''}{f'}-\frac{l(l+3)}{4}\frac{g''}{f'}$$
$$-\frac{l(l-1)^2(l+2)}{8}\left(\frac{f''}{f'}\right)^2-\frac{l(l-1)(2l+3)}{4}\frac{f''g'}{f'^2}-\frac{l(l+1)}{2}\left(\frac{g'}{f'}\right)^2,$$

$$f_l(0;0)=-\frac{\partial f}{\partial\varepsilon}-\frac{l(l-1)}{2}\frac{f''}{f'}-l\frac{g'}{f'},$$

and

$$g_l(0;0)=-\frac{\partial g}{\partial\varepsilon}-l\frac{\partial^2 f}{\partial x\partial\varepsilon}-\frac{l(l-1)(l-2)}{12}\frac{f'''}{f'}-\frac{l(l-1)}{4}\frac{g''}{f'}$$
$$-\frac{l(l-2)(l-1)^2}{8}\left(\frac{f''}{f'}\right)^2-\frac{l(l-1)(2l-3)}{4}\frac{f''g'}{f'^2}$$
$$-\frac{l(l-1)}{2}\left(\frac{g'}{f'}\right)^2-\frac{l(l-1)}{2}\frac{\partial f}{\partial\varepsilon}\frac{f''}{f'}-l\frac{\partial f}{\partial\varepsilon}\frac{g'}{f'}.$$

Thus we have $I=J$. Therefore $\det N\neq 0$ for sufficiently small $\varepsilon>0$. This shows that the problem (1.1) has a solution. The uniqueness of the solution follows from the Fredholm alternative theorem. The quantities $X=y'(-a)$ and $Y=y'(b)$ are obtained by solving (3.7). Substituting these into $v_l(x)$ and integrating $v_l(x)$, we get the solution $y(x)$. This completes the proof.

These results show that if (1.1a) is homogeneous the solution is almost zero within $(-a,b)$ and the slope of the solution at both endpoints is $O(\varepsilon^{-1})$. It should be noted that these slopes differ from those for the case when $l=0$. If (1.1a) is inhomogeneous, it follows that the solution is $O(\varepsilon^{-\max(2,l+1/2)+1})$ within $(-a,b)$.

Remark 1. The quantity J is related to the quantity σ defined by Cook and Eckhaus [3, p. 135] as follows:

$$J=-f'(0;0)\sigma.$$

One finds from this equation and the similar one for the case of $l=0$ that Theorem 1 and Theorem 2 are generalizations of their results on resonance, since they restrict their attention to the homogeneous problem (1.1) with $f(x;\varepsilon)$ and $g(x;\varepsilon)$ analytic functions of both x and ε.

Remark 2. When $f(x;\varepsilon)=-x$ and $g(x;\varepsilon)=g(x)$, we have

$$J=-(g')^2-\frac{2l+1}{2}g''.$$

This equation together with the similar one for the case of $l=0$ show that our Theorem 1 and Theorem 2 include the results of Lemma 3.1 by Kreiss and Parter [4, p. 244].

4. Examples. In this section we consider some illustrative examples. The first example is

$$\varepsilon y''-xy'+(x+c\varepsilon)y=0, \qquad -1\leqq x\leqq 1,$$
$$y(-1)=2, \qquad y(1)=-1,$$

where $c\neq -1$. Since all the conditions of Theorem 1 are satisfied, there exists a unique solution

$$y(x)=\begin{cases} 2\exp(x+1)\left(1-\dfrac{\int_{-1}^{x}\exp(-(1/2\varepsilon)(1-t^2))\,dt}{\int_{-1}^{0}\exp(-(1/2\varepsilon)(1-t^2))\,dt}\right)+O(\sqrt{\varepsilon}), & -1\leqq x\leqq 0,\\ -\exp(x-1)\left(1-\dfrac{\int_{x}^{1}\exp(-(1/2\varepsilon)(1-t^2))\,dt}{\int_{0}^{1}\exp(-(1/2\varepsilon)(1-t^2))\,dt}\right)+O(\sqrt{\varepsilon}), & 0\leqq x\leqq 1.\end{cases}$$

Next, to observe the influence of the right hand term on the nature of the solution, we analyze the problem

$$\varepsilon y''-xy'+(x+c\varepsilon)y=1, \qquad -1\leqq x\leqq 1,$$
$$y(-1)=2, \qquad y(1)=-1,$$

where $c\neq -1$. By an easy calculation, we get

$$\mu'_-(0)p_-(0)-\mu'_+(0)p_+(1)=\int_{-1}^{1}\exp\left(-\frac{t^2}{2\varepsilon}\right)dt+O(\varepsilon)$$

and

$$\alpha_-(0)+\alpha_+(1)=-(c+1)\int_{-1}^{1}\exp\left(-\frac{t^2}{2\varepsilon}\right)dt+O(\varepsilon).$$

Therefore, we have

$$\frac{p_-(x)}{\varepsilon}\left\{\mu_-(x)-\left(\mu_-(0)+\frac{\mu'_-(0)p_-(0)-\mu'_+(0)p_+(1)}{(\alpha_-(0)+\alpha_+(1))p_-(0)}\right)\frac{\lambda_-(x)}{\lambda_-(0)}\right\}$$
$$=\frac{\exp(x)}{(c+1)\varepsilon}\frac{\int_{-1}^{x}\exp(-(1-t^2)/2\varepsilon)\,dt}{\int_{-1}^{0}\exp(-(1-t^2)/2\varepsilon)\,dt}+O(\varepsilon^{-1/2})$$

for $-1<x\leqq 0$. Similarly, we have

$$\frac{p_+(1-x)}{\varepsilon}\left\{\mu_+(x)-\left(\mu_+(0)+\frac{\mu'_-(0)p_-(0)-\mu'_+(0)p_+(1)}{(\alpha_-(0)+\alpha_+(1))p_+(1)}\right)\frac{\lambda_+(x)}{\lambda_+(0)}\right\}$$
$$=\frac{\exp(x)}{(c+1)\varepsilon}\frac{\int_{x}^{1}\exp(-(1-t^2)/2\varepsilon)\,dt}{\int_{0}^{1}\exp(-(1-t^2)/2\varepsilon)\,dt}+O(\varepsilon^{-1/2})$$

for $0 \leqq x < 1$. This shows that the solution $y(x)$ is strongly influenced by the right hand term.

Finally, we solve the turning point problem

$$\varepsilon y'' - x(1+x^2)y' + (2+c\varepsilon)y = 0, \qquad -1 \leqq x \leqq 1,$$
$$y(-1) = -2, \qquad y(1) = 1$$

treated by Ackerberg and O'Malley [1]. The quantity J in Theorem 2 is given by $J = 12 - c$. Therefore, if $c \neq 12$, this problem has a unique solution. Let us choose $\delta > 0$ to satisfy $\delta^2(\delta^2+2) < 3/2$. Then, we have

$$y(x) = \begin{cases} -2p_-(x)\left(1 - \dfrac{\lambda_-(x)}{\lambda_-(-\delta)}\right) + O\left(\varepsilon^{-5/2} \exp\left(-\dfrac{\delta^2(\delta^2+2)}{4\varepsilon}\right)\right), & -1 \leqq x \leqq -\delta, \\ O\left(\varepsilon^{-5/2} \exp\left(-\dfrac{\delta^2(\delta^2+2)}{4\varepsilon}\right)\right), & -\delta \leqq x \leqq 0, \\ O\left(\varepsilon^{-5/2} \exp\left(-\dfrac{\delta^2(\delta^2+2)}{4\varepsilon}\right)\right), & 0 \leqq x \leqq \delta, \\ p_+(1-x)\left(1 - \dfrac{\lambda_+(x)}{\lambda_+(\delta)}\right) + O\left(\varepsilon^{-5/2} \exp\left(-\dfrac{\delta^2(\delta^2+2)}{4\varepsilon}\right)\right), & \delta \leqq x \leqq 1. \end{cases}$$

Acknowledgment. The author would like to express his deep gratitude to professor Seiiti Huzino in Kyushu University for his many stimulating suggestions and his very careful reading of this paper.

REFERENCES

[1] R. C. ACKERBERG AND R. E. O'MALLEY, JR., *Boundary layer problems exhibiting resonance*, Studies in Appl. Math., 49 (1970), pp. 277–295.

[2] R. A. CLARK, *Asymptotic solutions of a nonhomogeneous differential equation with a turning point*, Arch. Rational Mech. Anal., 12 (1963), pp. 34–51.

[3] L. P. COOK AND W. ECKHAUS, *Resonance in a boundary value problem of singular perturbation type*, Studies in Appl. Math., 52 (1973), pp. 129–139.

[4] H. O. KREISS AND S. V. PARTER, *Remarks on singular perturbations with turning points*, this Journal, 5 (1974), pp. 230–251.

[5] W. D. LAKIN, *Boundary value problems with a turning point*, Studies in Appl. Math., 51 (1972), pp. 261–275.

[6] B. J. MATKOWSKY, *On boundary layer problems exhibiting resonance*, SIAM Rev., 17 (1975), pp. 82–100.

[7] A. H. NAYFEH, *Perturbation Methods*, John Wiley, New York, 1973.

[8] R. E. O'MALLEY, JR., *On boundary value problems for a singularly perturbed differential equation with a turning point*, this Journal, 4 (1970), pp. 479–490.

[9] C. E. PEARSON, *On a differential equation of boundary layer type*, J. Math. and Phys., 47 (1968), pp. 134–154.

[10] L. B. RALL, *Computational Solution of Nonlinear Operator Equations*, John Wiley, New York, 1969.

[11] W. WASOW, *Asymptotic Expansions for Ordinary Differential Equations*, Interscience, New York, 1965.

[12] A. M. WATTS, *A singular perturbation problem with a turning point*, Bull. Austral. Math. Soc., 5 (1971), pp. 61–73.

[13] J. YARMISH, *Newton's method techniques for singular perturbations*, this Journal, 6 (1975), pp. 661–680.

[14] E. ZAUDERER, *Boundary value problems for a second order differential equation with a turning point*, Studies in Appl. Math., 51 (1972), pp. 411–413.

SIAM J. MATH. ANAL.
Vol. 9, No. 2, April 1978
Copyright © 1978 Society for Industrial and Applied Mathematics

ALGEBRAIC STRUCTURE AND FINITE DIMENSIONAL NONLINEAR ESTIMATION*

STEVEN I. MARCUS† AND ALAN S. WILLSKY‡

Abstract. The algebraic structure of certain classes of nonlinear systems is exploited in order to prove that the optimal estimators for these systems are recursive and finite dimensional. These systems are represented by certain Volterra series expansions or by bilinear systems with nilpotent Lie algebras. In addition, an example is presented, and the steady-state estimator for this example is discussed.

1. Introduction. Optimal recursive state estimators have been derived for very general classes of nonlinear stochastic systems [14], [7]. The optimal estimator requires, in general, an infinite dimensional computation to generate the conditional mean of the system state given the past observations. This computation involves either the solution of a stochastic partial differential equation for the conditional density or an infinite set of coupled ordinary stochastic differential equations for the conditional moments. However, the class of linear stochastic systems with linear observations and white Gaussian plant and observation noises has a particularly appealing structure, because the optimal state estimator consists of a finite dimensional linear system—the Kalman–Bucy filter [12].

In this paper we exploit the algebraic structure of certain other classes of systems, in order to prove that the optimal estimators for these systems are finite dimensional. The general class of systems is given by a linear Gauss–Markov process ξ which feeds forward into a nonlinear system with state x. Our goal is to estimate ξ and x given noisy linear observations of ξ. Specifically, consider the system

$$d\xi(t)=F(t)\xi(t)\,dt+G(t)\,dw(t), \tag{1.1}$$

$$dx(t)=a_0(x(t))\,dt+\sum_{i=1}^{N} a_i(x(t))\xi_i(t)\,dt, \tag{1.2}$$

$$dz(t)=H(t)\xi(t)\,dt+R^{1/2}(t)\,dv(t), \tag{1.3}$$

where $\xi(t)$ is an n-vector, $x(t)$ is a k-vector, $z(t)$ is a p-vector, w and v are independent standard Brownian motion processes, $R>0$, $\xi(0)$ is a Gaussian random variable independent of w and v, $x(0)$ is independent of $\xi(0)$, w, and v, and $\{a_i, i=0,\cdots,N\}$ are analytic functions of x. It will be assumed that $[F(t), G(t), H(t)]$ is completely controllable and observable. Also we define $Q(t)\triangleq G(t)G'(t)$.

The optimal estimate, with respect to a wide variety of criteria, of $x(t)$ given the observations $z^t\triangleq\{z(s), 0\leqq s\leqq t\}$, is the conditional mean $\hat{x}(t|t)$, also denoted by $E^t[x(t)]$ or $E[x(t)|z^t]$ [8] (henceforth we will freely interchange these three notations for the conditional expectation given the σ field $\sigma\{z(s), 0\leqq s\leqq t\}$ generated by the observation process up to time t). Thus our objective is the computation of $\hat{\xi}(t|t)$ and $\hat{x}(t|t)$. The computation of $\hat{\xi}(t|t)$ can be performed by the finite dimensional (linear)

* Received by the editors November 14, 1975.

† Department of Electrical Engineering, University of Texas at Austin, Austin, Texas 78712. The work of this author was supported in part by a National Science Foundation Fellowship, in part by the DoD Joint Services Electronics Program under AFSC Contract F44620-71-C-0091, and in part by the National Science Foundation under Grant ENG 76-11106.

‡ Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139. The work of this author was supported in part by the National Science Foundation under Grant GK-42090 and in part by the Air Force Office of Scientific Research under Grant 72-2273.

Kalman–Bucy filter; moreover, the conditional density of $\xi(t)$ given z^t is Gaussian with mean $\hat{\xi}(t|t)$ and nonrandom covariance $P(t)$ [12], [8]. However, the computation of $\hat{x}(t|t)$ requires in general an infinite dimensional system of equations. The purpose of this paper is to show that if $x(t)$ is characterized by a certain type of Volterra series expansion, or if $x(t)$ satisfies a certain type of bilinear equation, then $\hat{x}(t|t)$ can be computed with a *finite dimensional* nonlinear estimator.

This research is related to the recent work of Brockett [1]–[3] on algebraic and geometric methods in control theory and the work of Lo and Willsky [17], [25] on estimation for bilinear systems.

2. Volterra series and finite dimensional estimation. As shown by Brockett [2], [3] and d'Alessandro, Isidori and Ruberti [5] in the deterministic case, considerable insight can be gained by considering the Volterra series expansion of the system (1.2). The Volterra series expansion for the ith component of x is given by

$$x_i(t)=w_{0i}(t)+\sum_{j=1}^{\infty}\int_0^t\cdots\int_0^t\sum_{k_1,\cdots,k_j=1}^{n} w_{ji}^{(k_1,\cdots,k_j)}(t,\sigma_1,\cdots,\sigma_j)\cdot \xi_{k_1}(\sigma_1)\cdots\xi_{k_j}(\sigma_j)\,d\sigma_1\cdots d\sigma_j \tag{2.1}$$

where the jth *order kernel* $w_{ji}^{(k_1,\cdots,k_j)}$ is a locally bounded, piecewise continuous function. We will consider, without loss of generality [2], only *triangular kernels* which satisfy $w_{ji}^{(k_1,\cdots,k_j)}(t,\sigma_1,\cdots,\sigma_j)=0$ if $\sigma_{l+m}>\sigma_m$; $l,m=1,2,3,\cdots$. We say that a kernel $w(t,\sigma_1,\cdots,\sigma_j)$ is *separable* if it can be expressed as a finite sum

$$w(t,\sigma_1,\cdots,\sigma_j)=\sum_{i=1}^{m}\gamma_0^i(t)\gamma_1^i(\sigma_1)\gamma_2^i(\sigma_2)\cdots\gamma_j^i(\sigma_j). \tag{2.2}$$

Brockett [2] discusses the convergence of (2.1) in the deterministic case, but we will not consider this question in the general stochastic case. We will be more concerned with the case in which the linear-analytic system (1.2) has a finite Volterra series—that is, the expansion (2.1) has a finite number of terms. Brockett shows that a finite Volterra series has a bilinear realization if and only if the kernels are separable. Hence, a proof similar to that of Martin [20] of the existence and uniqueness of solutions to a bilinear system drived by the Gauss–Markov process (1.1) implies that a finite Volterra series in ξ with separable kernels is well defined in the mean-square sense.

With these preliminary concepts, the major results can be stated. The proofs are contained in this section and Appendix B; an example follows.

THEOREM 2.1. *Consider the linear system described by* (1.1) *and* (1.3), *and define the scalar-valued process*

$$x(t)=e^{\xi_j(t)}\eta(t) \tag{2.3}$$

where η is a finite Volterra series in ξ with separable kernels. Then $\hat{\eta}(t|t)$ and $\hat{x}(t|t)$ can be computed with a finite dimensional recursive system of nonlinear stochastic differential equations driven by the innovations $d\nu(t)\triangleq dz(t)-H(t)\hat{x}(t|t)\,dt$.

THEOREM 2.2. *Consider the linear system described by* (1.1) *and* (1.3), *and define the scalar-valued processes*

$$\eta(t)=\int_0^t\int_0^{\sigma_1}\cdots\int_0^{\sigma_{j-1}}\xi_{k_1}(\sigma_{m_1})\cdots\xi_{k_i}(\sigma_{m_i})\gamma_1(\sigma_1)\cdots\gamma_j(\sigma_j)\,d\sigma_1\cdots d\sigma_j, \tag{2.4}$$

$$x(t)=e^{\xi_l(t)}\eta(t) \tag{2.5}$$

where $\{\gamma_i\}$ *are deterministic functions of time and* $i > j$. *Then* $\hat{\eta}(t|t)$ *and* $\hat{x}(t|t)$ *can be computed with a finite dimensional recursive system of nonlinear stochastic differential equations driven by the innovations.*

The distinction between Theorems 2.1 and 2.2 lies in the fact that $i > j$ in (2.4)—i.e., there are more ξ_k's than integrals. On the other hand, each term in the finite Volterra series in (2.3) has $i = j$ and the σ_{m_k} are distinct. As Brockett [2] remarks, we can consider (2.4) as a single term in Volterra series if the kernel is allowed to contain impulse functions. As we will show in Lemma B.2, a term (2.4) with $i < j$ (more integrals than ξ_k's) can be rewritten as a Volterra term with $i = j$; so Theorem 2.1 also applies in this case.

The basic technique employed in the proofs of Theorems 2.1 and 2.2 is the augmentation of the state of the original system with the processes which are required in the nonlinear filtering equation (A.5)–(A.6) for $\hat{x}(t|t)$. For the classes of systems considered here, it is shown that only a finite number of additional states are required.

Proof of Theorem 2.1. We consider one term in the finite Volterra series; since the kernels are separable, we can assume without loss of generality that this term has the form

$$\eta(t) = \int_0^t \int_0^{\sigma_1} \cdots \int_0^{\sigma_{j-1}} \xi_{k_1}(\sigma_1) \cdots \xi_{k_j}(\sigma_j) \gamma_1(\sigma_1) \cdots \gamma_j(\sigma_j)\, d\sigma_1 \cdots d\sigma_j. \tag{2.6}$$

The theorem is proved by induction on j, the *order* of the Volterra term (2.6). We now give the proof for $j = 1$; the proof by induction is given in Appendix B. If $j = 1$, then

$$\eta(t) = \int_0^t \gamma_1(\sigma_1) \xi_{k_1}(\sigma_1)\, d\sigma_1 \tag{2.7}$$

and $\eta(t)$ is *linear* function of ξ. Hence, if the state ξ of (1.1) is augmented with η, the resulting system is also linear. Then the Kalman–Bucy filter for the system described by (1.1), (1.3), and (2.7) generates $\hat{\xi}(t|t)$ and $\hat{\eta}(t|t)$. In order to prove that $\hat{x}(t|t)$ is "finite dimensionally computable" (FDC), we need the following lemma. First we define, for $\sigma_1, \sigma_2 \leqq t$, the conditional cross-covariance matrix

$$P(\sigma_1, \sigma_2, t) = E[(\xi(\sigma_1) - \hat{\xi}(\sigma_1|t))(\xi(\sigma_2) - \hat{\xi}(\sigma_2|t))'|z^t] \tag{2.8}$$

(where $\hat{\xi}(\sigma|t) = E[\xi(\sigma)|z^t]$).

LEMMA 2.1. *The joint conditional density* $p_{\xi(\sigma_1),\xi(\sigma_2)}(\nu, \nu'|z^t)$ *is Gaussian with nonrandom conditional cross-covariance* $P(\sigma_1, \sigma_2, t)$*—i.e.,* $P(\sigma_1, \sigma_2, t)$ *is independent of* $\{z(s), 0 \leqq s \leqq t\}$.

Proof. First, the conditional density is Gaussian because ξ^t and z^t are jointly Gaussian random processes. Assume $\sigma_1 > \sigma_2$; then

$$p_{\xi(\sigma_1),\xi(\sigma_2)}(\nu, \nu'|z^t) = p_{\xi(\sigma_1)}(\nu|\xi(\sigma_2) = \nu', z^t) p_{\xi(\sigma_2)}(\nu'|z^t) \tag{2.9}$$

$$= p_{\xi(\sigma_1)}(\nu|\xi(\sigma_2) = \nu', z^t_{\sigma_2}) p_{\xi(\sigma_2)}(\nu'|z^t) \tag{2.10}$$

where $z^t_{\sigma_2} = \{z(s), \sigma_2 \leqq s \leqq t\}$.

Here (2.9) follows by the definition of the conditional density, and (2.10) is due to the Markov property of the process (ξ, z) [8]. Each of the densities in (2.10) is the result of a linear smoothing operation; hence, each is Gaussian with nonrandom covariance $P_{\sigma_1|\sigma_2}(t)$ and $P(\sigma_2, \sigma_2, t)$, respectively [16]. Also, for $\sigma > 0$, [11] $P(\sigma, \sigma, t) = [P^{-1}(\sigma) + P_B^{-1}(\sigma)]^{-1}$ where P_B is the error covariance of a Kalman filter running backward in time from t to σ, and $P_B^{-1}(t) \triangleq 0$. Due to the controllability of $[F, G]$, $P(\sigma)$

is invertible for all $\sigma > 0$ and $P_B(\sigma)$ is invertible for all $\sigma < t$ [28]; consequently, $P(\sigma, \sigma, t)$ is invertible for all $0 < \sigma \leqq t$. By the formula for the conditional covariance of a Gaussian distribution [8], we have for $0 \leqq \sigma_1 < \sigma_2 \leqq t$

$$P_{\sigma_1|\sigma_2}(t) = P(\sigma_1, \sigma_1, t) - P(\sigma_1, \sigma_2, t)P^{-1}(\sigma_2, \sigma_2, t)P'(\sigma_1, \sigma_2, t). \tag{2.11}$$

Since $P(\sigma_1, \sigma_2, t)$, $0 \leqq \sigma_1 < \sigma_2 < t$, can be computed from (2.11), it is also nonrandom; and since we have shown previously that $P(0, 0, t)$ is nonrandom, $P(\sigma_1, \sigma_2, t)$ is nonrandom for all $0 \leqq \sigma_1, \sigma_2 \leqq t$. □

This lemma allows the off-line computation of $P(\sigma_1, \sigma_2, t)$ via the equations of Kwakernaak [15] (for $\sigma_1 \leqq \sigma_2$)

$$\begin{aligned} P(\sigma_1, \sigma_2, t) &= P(\sigma_1)\Psi'(\sigma_2, \sigma_1) \\ &\quad - P(\sigma_1)\left[\int_{\sigma_2}^{t} \Psi'(\tau, \sigma_1)H'(\tau)R^{-1}(\tau)H(\tau)\Psi(\tau, \sigma_2)\, d\tau\right]P(\sigma_2), \end{aligned} \tag{2.12}$$

$$\frac{d}{dt}\Psi(t, \tau) = [F(t) - P(t)H'(t)R^{-1}(t)H(t)]\Psi(t, \tau); \ \Psi(\tau, \tau) = I \tag{2.13}$$

where the Kalman filter error covariance matrix $P(t) \triangleq P(t, t, t)$ is computed via the Riccati equation

$$\begin{aligned} \dot{P}(t) &= F(t)P(t) + P(t)F'(t) + Q(t) - P(t)H'(t)R^{-1}(t)H(t)P(t), \\ P(0) &= P_0. \end{aligned} \tag{2.14}$$

Recall [8] that the characteristic function of a Gaussian random vector y with mean m and covariance P is given by

$$M_y(u) = E[\exp(iu'y)] = \exp[iu'm - \tfrac{1}{2}u'Pu]. \tag{2.15}$$

Hence, by taking partial derivatives of the characteristic function (see Lemma B.1), we have

$$\begin{aligned} E^t[x(t)] &= \int_0^t \gamma_1(\sigma)E^t[e^{\xi_j(t)}\xi_{k_1}(\sigma)]\, d\sigma \\ &= \int_0^t \gamma_1(\sigma)[\hat{\xi}_{k_1}(\sigma|t) + P_{k_1,j}(\sigma, t, t)]\ e^{\hat{\xi}_j(t|t)+(1/2)P_{jj}(t)}\, d\sigma, \\ &= \left\{\int_0^t \gamma_1(\sigma)P_{k_1,j}(\sigma, t, t)\, d\sigma + E^t\left[\int_0^t \gamma_1(\sigma)\xi_{k_1}(\sigma)\, d\sigma\right]\right\} \cdot e^{\hat{\xi}_j(t|t)+(1/2)P_{jj}(t)} \\ &= \left\{\int_0^t \gamma_1(\sigma)P_{k_1,j}(\sigma, t, t)\, d\sigma + \hat{\eta}(t|t)\right\} e^{\hat{\xi}_j(t|t)+(1/2)P_{jj}(t)}. \end{aligned} \tag{2.16}$$

Since the first term in (2.16) is nonrandom and $\hat{\eta}(t|t)$ and $\hat{\xi}(t|t)$ can be computed with a Kalman–Bucy filter, $\hat{x}(t|t)$ is indeed FDC for the case $j = 1$.

The induction step of the proof of Theorem 2.1 is given in Appendix B. A crucial component of the proof is Lemma B.1, which expresses higher order moments of a Gaussian distribution in terms of the lower moments. Notice that in equation (2.16) we have interchanged the operations of integration and conditional expectation. This is justified by the version of the Fubini theorem proved in [18]; since we will be dealing only with integrals of products of Gaussian random processes, the use of the Fubini theorem is easily justified, and we will use it without further comment.

The proof of Theorem 2.2 is almost identical to that of Theorem 2.1; the differences are explained in Appendix B. We now present an example to illustrate the basic concepts of these theorems; this example is a special case of Theorem 2.2. However, we will need one preliminary lemma.

LEMMA 2.2. *The conditional cross-covariance satisfies*

$$P(\sigma, t, t) = K(t, \sigma)P(t) \tag{2.17}$$

where

$$\frac{d}{dt}K'(t, \sigma) = -[F'(t) + P^{-1}(t)Q(t)]K'(t, \sigma); \; K'(\sigma, \sigma) = I. \tag{2.18}$$

Proof. Let

$$\tilde{P}(\sigma, t) \triangleq E[(\xi(\sigma) - \hat{\xi}(\sigma|\sigma))(\xi(t) - \hat{\xi}(t|t))']$$

and consider

$$P(\sigma, t, t) - \tilde{P}(\sigma, t) = E[(\hat{\xi}(\sigma|\sigma) - \hat{\xi}(\sigma|t))(\xi(t) - \hat{\xi}(t|t))'|z^t].$$

Since $\hat{\xi}(\sigma|\sigma) - \hat{\xi}(\sigma|t)$ is measurable with respect to the σ-field $\sigma(z^t)$, the projection theorem [22] implies that $P(\sigma, t, t) - \tilde{P}(\sigma, t) = 0$. The proof is concluded by noting that [11] $\tilde{P}(\sigma, t) = K(t, \sigma)P(t)$. □

Example 2.1. Consider the system described by

$$\begin{bmatrix} d\xi_1(t) \\ d\xi_2(t) \end{bmatrix} = \begin{bmatrix} -\alpha & 0 \\ 0 & -\beta \end{bmatrix} \begin{bmatrix} \xi_1(t) \\ \xi_2(t) \end{bmatrix} dt + \begin{bmatrix} dw_1(t) \\ dw_2(t) \end{bmatrix}, \tag{2.19}$$

$$dx(t) = (-\gamma x(t) + \xi_1(t)\xi_2(t))\, dt, \tag{2.20}$$

$$\begin{bmatrix} dz_1(t) \\ dz_2(t) \end{bmatrix} = \begin{bmatrix} \xi_1(t) \\ \xi_2(t) \end{bmatrix} dt + \begin{bmatrix} dv_1(t) \\ dv_2(t) \end{bmatrix} \tag{2.21}$$

where $\alpha, \beta, \gamma > 0$, w_1, w_2, v_1, and v_2 are independent, zero mean, unit variance Wiener processes, $\xi_1(0)$ and $\xi_2(0)$ are independent Gaussian random variables which are also independent of the noise processes, and $x(0) = 0$.

The conditional expectation $\hat{x}(t|t)$ satisfies the nonlinear filtering equation (A.5)–(A.6):

$$\begin{aligned} d\hat{x}(t|t) = E^t[-\gamma x(t) + \xi_1(t)\xi_2(t)]\, dt + \Big\{ & E^t\Big[\int_0^t e^{-\gamma(t-s)}\xi_1(s)\xi_2(s)\, ds \cdot \xi'(t)\Big] \\ & - E^t\Big[\int_0^t e^{-\gamma(t-s)}\xi_1(s)\xi_2(s)\, ds\Big]\hat{\xi}'(t|t)\Big\}\, d\nu(t) \end{aligned} \tag{2.22}$$

where $\xi(t) = [\xi_1(t), \xi_2(t)]'$ and the innovations process ν is given by

$$d\nu(t) = dz(t) - \hat{\xi}(t|t)\, dt. \tag{2.23}$$

Recall that the conditional covariance $P(t)$ of $\xi(t)$ given z^t satisfies the Riccati equation (2.14). Since $\xi_1(0)$ and $\xi_2(0)$ are independent, it is not difficult to show that $P_{12}(t) = P_{21}(t) = 0$ for all t. From (2.17)–(2.18) we can compute

$$P(\sigma, t, t) = \begin{bmatrix} P_{11}(t)\exp[\alpha(t-\sigma) - \int_\sigma^t P_{11}^{-1}(s)\, ds] & 0 \\ 0 & P_{22}(t)\exp[\beta(t-\sigma) - \int_\sigma^t P_{22}^{-1}(s)\, ds] \end{bmatrix}. \tag{2.24}$$

These facts and equation (B.3a) imply that the transpose of the *gain term* in (2.22) is

$$E^t\left[\int_0^t e^{-\gamma(t-s)}\xi_1(s)\xi_2(s)\xi(t)\,ds\right]-E^t\left[\int_0^t e^{-\gamma(t-s)}\xi_1(s)\xi_2(s)\,ds\right]\hat{\xi}(t|t)$$

$$=\int_0^t e^{-\gamma(t-s)}(E^t[\xi_1(s)\xi_2(s)\xi(t)]-E^t[\xi_1(s)\xi_2(s)]E^t[\xi(t)])\,ds$$

(2.25a) $$=E^t\left\{\int_0^t e^{-\gamma(t-s)}\begin{bmatrix}0 & P_{11}(s,t,t)\\ P_{22}(s,t,t) & 0\end{bmatrix}\begin{bmatrix}\xi_1(s)\\ \xi_2(s)\end{bmatrix}ds\right\}$$

(2.25b) $$=E^t\begin{bmatrix}\eta_1(t)P_{11}(t)\\ \eta_2(t)P_{22}(t)\end{bmatrix}$$

where

(2.26) $$\begin{bmatrix}\dot\eta_1(t)\\ \dot\eta_2(t)\end{bmatrix}=\begin{bmatrix}\alpha-\gamma-P_{11}^{-1}(t) & 0\\ 0 & \beta-\gamma-P_{22}^{-1}(t)\end{bmatrix}\begin{bmatrix}\eta_1(t)\\ \eta_2(t)\end{bmatrix}+\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}\xi_1(t)\\ \xi_2(t)\end{bmatrix},$$
$$\eta_1(0)=\eta_2(0)=0.$$

In other words, the argument of the conditional expectation in (2.25a) can be realized as the output of a finite dimensional linear system with state $\eta(t)=[\eta_1(t),\ \eta_2(t)]'$ satisfying (2.26).

Thus the finite dimensional optimal estimator for the system (2.19)–(2.21) is constructed as follows (see Fig. 1). First we augment the state ξ of (2.19) with the state η of (2.26). Then the Kalman–Bucy filter for the linear system (2.19), (2.26), with observations (2.21), computes the conditional expectations $\hat{\xi}(t|t)$ and $\hat{\eta}(t|t)$. Finally,

(2.27) $$d\hat{x}(t|t)=[-\gamma\hat{x}(t|t)+\hat{\xi}_1(t|t)\hat{\xi}_2(t|t)]\,dt+\hat{\eta}'(t|t)P(t)\,d\nu(t),$$
$$\hat{x}(0|0)=0.$$

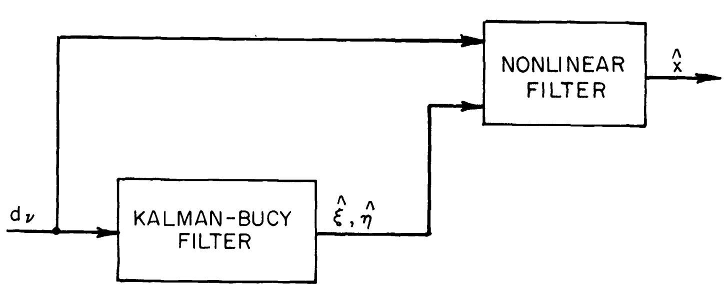

FIG. 1. *Block diagram of the optimal filter for Example* 2.1.

We now discuss the steady-state behavior of the optimal filter. Since the linear system (2.19) is asymptotically stable (and hence detectable) and controllable, the Riccati equation (2.14) has a unique positive-definite steady-state solution P [28]; a simple computation shows that

(2.28) $$P=\begin{bmatrix}P_{11} & 0\\ 0 & P_{22}\end{bmatrix}=\begin{bmatrix}-\alpha+\sqrt{\alpha^2+1} & 0\\ 0 & -\beta+\sqrt{\beta^2+1}\end{bmatrix}.$$

Thus, in steady-state, the augmented linear system (2.19), (2.26) is time-invariant. Now consider the eigenvalues of (2.26) in steady-state:

$$\alpha-\gamma-P_{11}^{-1}=\alpha-\gamma-(-\alpha+\sqrt{\alpha^2+1})^{-1}=-\gamma-\sqrt{\alpha^2+1},$$
$$\beta-\gamma-P_{22}^{-1}=\beta-\gamma-(-\beta+\sqrt{\beta^2+1})^{-1}=-\gamma-\sqrt{\beta^2+1}.$$

Consequently, the augmented linear system is also asymptotically stable and controllable in steady-state. Let the conditional covariance matrix of the augmented state $[\xi(t), \eta(t)]$ given z^t be denoted by $S(t)$. Then the Riccati equation satisfied by $S(t)$ has a unique positive-definite steady-state solution S (notice that $S_{11}=P_{11}$ and $S_{22}=P_{22}$).

The steady-state Kalman–Bucy filter [8] for the augmented system (2.19), (2.26) is easily computed to be

$$\begin{bmatrix} d\hat{\xi}_1(t|t) \\ d\hat{\xi}_2(t|t) \\ d\hat{\eta}_1(t|t) \\ d\hat{\eta}_2(t|t)\end{bmatrix}=\begin{bmatrix} -\alpha & 0 & 0 & 0 \\ 0 & -\beta & 0 & 0 \\ 0 & 1 & -\gamma-\sqrt{\alpha^2+1} & 0 \\ 1 & 0 & 0 & -\gamma-\sqrt{\beta^2+1}\end{bmatrix}\begin{bmatrix}\hat{\xi}_1(t|t) \\ \hat{\xi}_2(t|t) \\ \hat{\eta}_1(t|t) \\ \hat{\eta}_2(t|t)\end{bmatrix}dt$$
$$(2.29)\qquad +\begin{bmatrix} P_{11} & 0 \\ 0 & P_{22} \\ 0 & S_{23} \\ S_{14} & 0\end{bmatrix}\begin{bmatrix} d\nu_1(t) \\ d\nu_2(t)\end{bmatrix}$$

where

$$S_{14}=\frac{P_{11}P_{22}}{P_{11}P_{22}+(\alpha-\beta+\gamma)P_{22}+1},\qquad S_{23}=\frac{P_{11}P_{22}}{P_{11}P_{22}+(\beta-\alpha+\gamma)P_{11}+1}$$

(here P_{11} and P_{22} are defined in (2.28)). The conditional expectation $\hat{x}(t|t)$ is computed according to

$$(2.30)\qquad \begin{aligned} d\hat{x}(t|t)&=[-\gamma\hat{x}(t|t)+\hat{\xi}_1(t|t)\hat{\xi}_2(t|t)]\,dt+\hat{\eta}'(t|t)\,P\,d\nu(t),\\ \hat{x}(0|0)&=0\end{aligned}$$

which is a nonlinear, time-invariant equation.

We note that the stability of the original linear system is not necessary for the existence of the steady-state optimal filter in this example; in fact, a weaker sufficient condition is the detectability [28] of the linear system (2.19), (2.21) and the positivity of γ in (2.20). The generalization of this result to other systems is presently being investigated.

3. Finite dimensional estimators for bilinear systems. In this section the results of the previous section are applied, with the aid of some concepts from the theory of Lie algebra [23], to prove that the optimal estimators for certain bilinear systems are finite dimensional. Consider the system described by (1.1), (1.3), and the bilinear system [1], [10]

$$(3.1)\qquad \dot{X}(t)=\left(A_0+\sum_{i=1}^{N}\xi_i(t)A_i\right)X(t);\qquad X(0)=I$$

where X is a $k\times k$ matrix. We associate with (3.1) the Lie algebra $\mathscr{L}\triangleq\{A_0, A_1, \ldots, A_N\}_{LA}$, the smallest Lie algebra containing $A_0, A_1, \ldots, A_N$; the ideal $\mathscr{L}_0$ in $\mathscr{L}$ generated by $\{A_1, \ldots, A_N\}$; the group $G\triangleq\{\exp\mathscr{L}\}_G$, the smallest group generated by $\{\exp A\}$ for all $A\in\mathscr{L}$; and the subgroup $G_0\triangleq\{\exp\mathscr{L}_0\}_G$ [10], [18], [19], [26], [27].

DEFINITION 3.1 [23]. A Lie algebra $\mathscr{L}$ is *solvable* if the *derived series* of ideals

$$\mathscr{L}^{(0)}=\mathscr{L},$$
$$\mathscr{L}^{(n+1)}=[\mathscr{L}^{(n)},\mathscr{L}^{(n)}]=\{[A,B]|A,B\in\mathscr{L}^{(n)}\},\qquad n\geqq 0,$$

is the trivial ideal $\{0\}$ for some n. $\mathscr{L}$ is *nilpotent* if the *lower central* series of ideals

$$\mathscr{L}^{0}=\mathscr{L},$$
$$\mathscr{L}^{n+1}=[\mathscr{L},\mathscr{L}^{n}]=\{[A,B]|A\in\mathscr{L}^{n}\},\qquad n\geqq 0,$$

is $\{0\}$ for some n. $\mathscr{L}$ is Abelian if $\mathscr{L}^{(1)}=\mathscr{L}^1=\{0\}$. Note that Abelian$\Rightarrow$nilpotent$\Rightarrow$ solvable, but none of the reverse implications hold in general.

A useful structural result for nilpotent Lie algebras is presented in the following lemma [23, p. 224].

LEMMA 3.1. *A matrix Lie algebra $\mathscr{L}$ is nilpotent if and only if there exists a (possibly complex-valued) nonsingular matrix P such that, for all $A\in\mathscr{L}$, PAP^{-1} has the block diagonal form*

$$\begin{bmatrix} \begin{bmatrix}\Phi_1(A) & & *\\ & \ddots & \\ 0 & & \Phi_1(A)\end{bmatrix} & & 0\\ & \begin{bmatrix}\Phi_2(A) & & *\\ & \ddots & \\ & & \Phi_2(A)\end{bmatrix} & \\ 0 & & \ddots \end{bmatrix} \tag{3.2}$$

(this will be called the nilpotent canonical form*). The functions $\Phi_k\colon\mathscr{L}\to C$ are linear. Furthermore, $\Phi_k([\mathscr{L},\mathscr{L}])=\{0\}$.*

It is easy to show, using Brockett's results [2] on finite Volterra series, that each term in (2.3) can be realized by a bilinear system of the form

$$\dot{x}(t)=\xi_j(t)x(t)+\sum_{k=1}^{n}A_k(t)\xi_k(t)x(t) \tag{3.3}$$

where x is a k-vector and the A_j are strictly upper triangular (zero on and below the main diagonal). For such systems, the Lie algebra $\mathscr{L}_0$ is nilpotent. In this section we will show conversely that if the Lie algebra $\mathscr{L}_0$ corresponding to the bilinear system (3.1) is nilpotent, then each component of the solution to (3.1) can be written as a finite sum of terms of the form (2.3). Hence, such systems also have finite dimensional estimators; this result is summarized in the next theorem.

THEOREM 3.1. *Consider the system described by* (1.1), (1.3), *and* (3.1) *and assume that $\mathscr{L}_0$ is a nilpotent Lie algebra. Then the conditional expectation $\hat{X}(t|t)$ can be computed with a finite dimensional system of nonlinear differential equations driven by the innovations.*

Remarks. (i) It can easily be shown that if $\mathscr{L}_0$ is nilpotent, then $\mathscr{L}$ is solvable; however, the converse is not true. Hence, $\mathscr{L}$ is always solvable in Theorem 3.1.

(ii) Theorem 3.1 provides a generalization of the work of Lo and Willsky [17] (in which $\mathscr{L}$ is Abelian) and Willsky [25]. The Abelian discrete-time problem is also considered by Johnson and Stear [9].

(iii) The model considered in Theorem 3.1 is motivated by a problem in strapdown inertial navigation [18], [26]. However, in the navigation problem $\mathscr{L}_0$ is not nilpotent (in fact, $\mathscr{L} = so(3)$ is simple [23]), so Theorem 3.1 does not apply.

(iv) Using the notation of Brockett [2], it is easily seen that the pth order moments $X^{[p]}(t)$ satisfy an equation of the form (3.1) (with different coefficient matrices $A_{i_{[p]}}$), and hence $X^{[p]}(t|t)$ can also be computed with a finite dimensional estimator. In particular, the performance of the estimator of Theorem 3.1 can be evaluated by computing the conditional covariance of $X(t)$ given z^t in this manner.

Theorem 3.1 is proved via a series of lemmas which reduce the estimation problem to the case in which $\mathscr{L}$ is a particular nilpotent Lie algebra. The first lemma generalizes a result of Willsky [25], Brockett [1], and Krener [13] (the proof is analogous and will be omitted).

LEMMA 3.2. *Consider the system described by* (1.1), (1.3), *and* (3.1) *and define the* $k \times k$ *matrix-valued process*

$$Y(t) = e^{-A_0 t} X(t). \tag{3.4}$$

Then there exists a deterministic matrix-valued function $D(t)$ *such that* Y *satisfies*

$$\dot{Y}(t) = \left[\sum_{i=1}^{M} H_i y_i(t)\right] Y(t); \qquad Y(0) = I \tag{3.5}$$

where $\{H_1, \cdots, H_M\}$ *is a basis for* $\mathscr{L}_0$ *and*

$$y(t) = D(t)\xi(t). \tag{3.6}$$

In addition, $\hat{X}$ *can be computed according to*

$$\hat{X}(t|t) = e^{A_0 t}\hat{Y}(t|t). \tag{3.7}$$

Lemma 3.2 enables us, without loss of generality, to examine the estimation problem for $Y(t)$ evolving on the subgroup $G_0 = \{\exp \mathscr{L}_0\}_G$, rather than for $X(t)$ evolving on the full Lie group $G = \{\exp \mathscr{L}\}_G$. Thus, we need only consider the case in which $A_0 = 0$ and $\mathscr{L} = \mathscr{L}_0$ is nilpotent in order to prove Theorem 3.1.

By means of Lemma 3.1 the problem can be further reduced to the consideration of Lie algebras in nilpotent canonical form.

LEMMA 3.3. *Consider the system described by* (1.1), (1.3), *and* (3.1), *where* $A_0 = 0$ *and* $\mathscr{L}$ *is nilpotent. Then there exists a* (*possibly complex-valued*) *nonsingular matrix* P *such that*

$$\hat{X}(t|t) = P^{-1}\hat{Y}(t|t)P \tag{3.8}$$

where Y *satisfies* (3.5) *and* $[\{H_1, \cdots, H_M\}$ *are in nilpotent canonical form.*

Proof. According to Lemma 3.1, there exists a nonsingular matrix P such that $P\mathscr{L}P^{-1}$ is in nilpotent canonical form. If we define $H_i = PA_iP^{-1}$, then $X(t) = P^{-1}Y(t)P$, where Y satisfies (3.5). Hence, $\hat{X}(t|t) = P\hat{Y}(t|t)P^{-1}$ and the lemma is proved. □

Finally, by means of the following trivial lemma, we reduce the problem to the consideration of one block in the nilpotent canonical form.

LEMMA 3.4. *Consider the system described by* (1.1), (1.3), *and* (3.1), *where* $A_0 = 0$ *and* $\{A_1, \cdots, A_N\}$ *are in nilpotent canonical form. Then* $X(t)$ *has a block diagonal form conformable with that of* $\{A_1, \cdots, A_N\}$.

Let $gn(m)$ denote the Lie algebra of upper triangular $m \times m$ matrices with equal diagonal elements. Then Lemma 3.4 implies that the bilinear system (3.1) can be viewed as the "direct sum" of a number of decoupled k_j-dimensional subsystems

$$\dot{X}^j(t)=\left[\sum_{i=1}^{N}\xi_i(t)A_i^j\right]X^j(t);\qquad X^j(0)=I \tag{3.9}$$

where $A_1^j,\cdots,A_N^j$ belong to $gn(k_j)$. Hence, Theorem 3.1 will be established when we prove the following lemma.

LEMMA 3.5. *Consider the system described by* (1.1), (1.3), *and* (3.1), *where* $A_0=0$ *and* $\{A_1,\cdots,A_N\}\in gn(k)$. *Then each element of the solution* $X(t)$ *of* (3.1) *can be expressed in the form*

$$\exp\left(\sum_{i=1}^{N}\alpha_i\int_0^t\xi_i(s)\,ds\right)\eta(t) \tag{3.10}$$

where η *is a finite Volterra series in* ξ *with separable kernels. Hence, Theorem* 2.1 *implies that* $\hat{X}(t|t)$ *can be computed with a finite dimensional system of nonlinear stochastic differential equations.*

Proof. Since $\{A_1,\cdots,A_N\}\in gn(k)$, the bilinear equation (3.1) can be rewritten in the form

$$\dot{X}(t)=\left[\left(\sum_{i=1}^{N}\alpha_i\xi_i(t)\right)I+\sum_{i=1}^{N}\xi_i(t)B_i\right]X(t) \tag{3.11}$$

where α_i are constants, I denotes the $k\times k$ identity matrix, and $B_1,\cdots,B_N$ are *strictly* upper triangular (zero on the diagonal). It is easy to show that

$$X(t)=\exp\left(\sum_{i=1}^{N}\alpha_i\int_0^t\xi_i(s)\,ds\right)Y(t)$$

where Y satisfies

$$\dot{Y}(t)=\left[\sum_{i=1}^{N}\xi_i(t)B_i\right]Y(t);\qquad Y(0)=I. \tag{3.12}$$

Since the $\{B_i\}$ are strictly upper triangular, the solution of (3.12) can be written as a finite Peano–Baker (Volterra) series [2], and each element of $X(t)$ can be expressed in the form (3.10). □

4. Conclusions. It is shown in [18] that if $\mathscr{L}_0$ is not nilpotent, then the optimal estimator for (1.1), (1.3), and (3.1) is infinite dimensional. Thus, the results of this paper cannot be generalized to much larger classes of systems.

However, the papers of Fliess [6] and Sussmann [24] show that, in the deterministic case with bounded inputs, any causal and continuous input-output map on a finite interval can be uniformly approximated by a bilinear system of the form (3.1) in which $A_0, A_1,\cdots,A_N$ are all strictly upper triangular. For such a bilinear system both $\mathscr{L}_0$ and $\mathscr{L}$ are nilpotent Lie algebras. Stochastic analogues of this result are currently being investigated. The implication of such a result would be that suboptimal estimators for a large class of nonlinear stochastic systems could be constructed using the results of this paper.

Appendix A. General nonlinear filtering equations. In this Appendix we state some results on nonlinear filtering [7], [8], [14]. Consider a model in which the state evolves according to the Ito stochastic differential equation

$$dx(t)=f(x(t),t)\,dt+G(x(t),t)\,dw(t) \tag{A.1}$$

and the observed process is the solution of the vector Ito equation

$$dz(t)=h(x(t),t)\,dt+R^{1/2}(t)\,dv(t). \tag{A.2}$$

Here $x(t)$ is an n-vector, $z(t)$ is a p-vector, $R^{1/2}$ is the unique positive definite square root of the positive definite matrix R, and v and w are independent Brownian motion (Wiener) processes such that

$$E[w(t)w'(s)]=\int_0^{\min(t,s)} Q(\tau)\,d\tau, \tag{A.3}$$

$$E[v(t)v'(s)]=\min(t,s)\cdot I. \tag{A.4}$$

For any integrable random process $\alpha(t)$, we denote $E(\alpha(t)|z(s),\ 0\leqq s\leqq t)$ by $\hat{\alpha}(t|t)$ or $E^t[\alpha(t)]$. Then, [7], [8], [14], the conditional mean $\hat{x}(t|t)$ satisfies

$$d\hat{x}(t|t)=E^t[f(x(t),t)]\,dt+\{E^t[x(t)h'(x(t),t)]-\hat{x}(t|t)E^t[h'(x(t),t)]\}R^{-1}(t)\,d\nu(t) \tag{A.5}$$

where the *innovations process* ν is defined by

$$d\nu(t)=dz(t)-E^t[h(x(t),t)]\,dt. \tag{A.6}$$

Appendix B. Proofs of Theorems 2.1 and 2.2.

B.1. Preliminary results. In this section we present some preliminary results which are crucial in the proofs of Theorems 2.1 and 2.2. The first lemma follows easily from some identities of Miller [21].

LEMMA B.1. *Let $x=[x_1,\cdots,x_k]'$ be a Gaussian random vector with mean m, covariance matrix P, and characteristic function M_x. Then, if $l\leqq k$,*

$$\begin{aligned}\frac{\partial^l}{\partial u_1\cdots\partial u_l}M_x(u_1,\cdots,u_k)=\Big\{&\varepsilon_1\cdots\varepsilon_l-\sum P_{j_1j_2}\varepsilon_{j_3}\cdots\varepsilon_{j_l}\\&+\sum P_{j_1j_2}P_{j_3j_4}\varepsilon_{j_5}\cdots\varepsilon_{j_l}-\cdots\Big\}M_x(u_1,\cdots,u_k)\end{aligned} \tag{B.1}$$

where

$$\varepsilon_j=im_j-\sum_{n=1}^{k}u_nP_{jn} \tag{B.2}$$

and the sums in (B.1) are over all possible combinations of pairs of the $\{j_i,\ i=1,\cdots,l\}$. Also,

$$E[x_1x_2\cdots x_k]=E[x_k]E[x_1x_2\cdots x_{k-1}]+\sum_{j_1=1}^{k-1}P_{kj_1}E[x_{j_2}x_{j_3}\cdots x_{j_{k-1}}] \tag{B.3a}$$

$$\begin{aligned}&=E[x_1\cdots x_i]E[x_{i+1}\cdots x_k]+\sum P_{j_1l_{i+1}}E[x_{j_2}\cdots x_{j_i}]E[x_{l_{i+2}}\cdots x_{l_k}]\\&\quad+\sum P_{j_1l_{i+1}}P_{j_2l_{i+2}}E[x_{j_3}\cdots x_{j_i}]E[x_{l_{i+3}}\cdots x_{l_k}]+\cdots\end{aligned} \tag{B.3b}$$

$$\begin{aligned}&=m_1\cdots m_k+\sum P_{j_1j_2}m_{j_3}\cdots m_{j_k}\\&\quad+\sum P_{j_1j_2}P_{j_3j_4}m_{j_5}\cdots m_{j_k}+\cdots\end{aligned} \tag{B.3c}$$

where the sums in (B.3b, c) are defined as in (B.1); also, in (B.3b), $\{j_\alpha,\ \alpha=1,\cdots,i\}$ is a permutation of $\{1,\cdots,i\}$ and $\{l_\alpha,\ \alpha=i+1,\cdots,k\}$ is a permutation of $\{i+1,\cdots,k\}$.

In the remainder of this Appendix it will be assumed that ξ and z are Gauss–Markov processes satisfying (1.1) and (1.3), respectively. We now define classes of random processes which occur as the jth order term in a Volterra series expansion in ξ

with separable kernels, and we prove some lemmas relating these to other relevant processes.

DEFINITION B.1. The space Λ_j of *Volterra terms of order* j is the vector space over R consisting of all scalar-valued random processes λ_j of the form

$$\lambda_j(t)=\sum_{i=1}^{N}\gamma_0^i(t)\lambda_j^i(t) \tag{B.4}$$

where

$$\lambda_j^i(t)=\int_0^t\int_0^{\sigma_1}\cdots\int_0^{\sigma_{j-1}}\gamma_1^i(\sigma_1)\cdots\gamma_j^i(\sigma_j)\xi_{k_{1,i}}(\sigma_1)\cdots\xi_{k_{j,i}}(\sigma_j)\,d\sigma_1\cdots d\sigma_j \tag{B.5}$$

where for each i, $\{\xi_{k_{1,i}},\cdots,\xi_{k_{j,i}}\}$ are not necessarily distinct elements of ξ, and $\{\gamma_l^i\}$ are locally bounded, piecewise continuous, deterministic functions of time. We denote by $\hat{\Lambda}_j$ the space of all processes

$$\hat{\lambda}_j(t|t)\triangleq E[\lambda_j(t)|z^t],\qquad \text{where } \lambda_j\in\Lambda_j.$$

The next lemma, which is due to Brockett [4], shows that terms of the form (2.4) with $i<j$ (more integrals than ξ_k's) are in fact elements of Λ_i.

LEMMA B.2. *Let ξ satisfy* (1.1) *and consider the scalar-valued process*

$$\eta(t)=\int_0^t\int_0^{\sigma_1}\cdots\int_0^{\sigma_{j-1}}\gamma_1(\sigma_1)\cdots\gamma_j(\sigma_j)\xi_{k_1}(\sigma_{m_1})\cdots\xi_{k_i}(\sigma_{m_i})\,d\sigma_1\cdots d\sigma_j \tag{B.6}$$

where γ_i are as in Definition B.1, *$m_n\neq m_l$ for $n\neq l$, and $i<j$. Then $\eta\in\Lambda_i$.*

Proof. It is easy to show using the construction of Brockett [2, Thm. 4] that $\eta(t)$ has a realization as a time-varying bilinear system

$$\dot{x}(t)=A(t)x(t)+\sum_{l=1}^{i}\xi_{k_l}(t)B_l(t)x(t), \tag{B.7}$$

$$\eta(t)=x_1(t) \tag{B.8}$$

where $A(t)$ and $\{B_l(t)\}$ are strictly upper triangular matrices. The Volterra series for (B.7) can be expressed via the Peano–Baker series [2], and the Volterra series is finite because $A(t)$ and $\{B_l(t)\}$ are upper triangular. In fact, because the original expression (B.6) contains only the product of i components of ξ, the Volterra expansion of $\eta(t)=x_1(t)$ will contain *only* an ith order term

$$\eta(t)=\int_0^t\int_0^{\sigma_1}\cdots\int_0^{\sigma_{i-1}}\left[\sum_{l=1}^{m}\gamma_1^l(\sigma_1)\cdots\gamma_i^l(\sigma_i)\right]\xi_{n_1}(\sigma_1)\cdots\xi_{n_i}(\sigma_i)\,d\sigma_1\cdots d\sigma_i \tag{B.9}$$

where $\{n_l, l=1,\cdots,i\}$ is a permutation of the $\{k_l, l=1,\cdots,i\}$ of (B.6). Hence $\eta\in\Lambda_i$. □

Recall that the conditional cross-covariance $P(\sigma_1,\sigma_2,t)$ (defined in (2.8)) was shown to be nonrandom in Lemma 2.1; it can be computed from Kwakernaak's equations (2.12)–(2.14). The following lemma shows that $P_{ij}(\sigma_1,\sigma_2,t)$ is a separable kernel.

LEMMA B.3. *$P_{ij}(\sigma_1,\sigma_2,$ t) is a separable kernel; i.e., it can be expressed in the form*

$$P_{ij}(\sigma_1,\sigma_2,t)=\sum_{k=1}^{m}\gamma_0^k(t)\gamma_1^k(\sigma_1)\gamma_2^k(\sigma_2). \tag{B.10}$$

Proof. Assume $\sigma_1 \leqq \sigma_2 \leqq t$. Then it follows from (2.12) that, for arbitrary real numbers α, β, and δ,

$$P(\sigma_1, \sigma_2, t) = P(\sigma_1)\Psi'(\alpha, \sigma_1)\Big[\Psi'(\sigma_2, \alpha) - \int_{\sigma_2}^{\beta} \Psi'(\tau, \alpha)H'(\tau)R^{-1}(\tau)H(\tau)\Psi(\tau, \sigma_2)\, d\tau \cdot P(\sigma_2)$$

$$- \int_{\beta}^{t} \Psi'(\tau, \alpha)H'(\tau)R^{-1}(\tau)H(\tau)\Psi(\tau, \delta)\, d\tau \cdot \Psi(\delta, \sigma_2)P(\sigma_2)\Big] \tag{B.11}$$

$$\triangleq A(\sigma_1)[B(\sigma_2) + C(t)D(\sigma_2)].$$

Hence, if e_i denotes the ith unit vector in R^n, it is obvious from (B.11) that

$$P_{ij}(\sigma_1, \sigma_2, t) = e_i' P(\sigma_1, \sigma_2, t) e_j \tag{B.12}$$

has the form (B.10) for some functions $\{\gamma_l^m(t)\}$. □

The next lemma proves that certain processes which occur in the proof of Theorem 2.1 are elements of Λ_j.

LEMMA B.4. *Let ξ satisfy* (1.1), *and consider the scalar-valued process*

$$\eta(t) = \int_0^t \int_0^{\sigma_1} \int_0^{\sigma_{j-1}} P_{n_1 n_2}(\sigma_{m_1}, \sigma_{m_2}, t) \cdots P_{n_{l-1} n_l}(\sigma_{m_{l-1}}, \sigma_{m_l}, t)$$

$$\cdot \gamma_1(\sigma_1) \cdots \gamma_j(\sigma_j) \xi_{k_1}(\sigma_1) \cdots \xi_{k_j}(\sigma_j)\, d\sigma_1 \cdots d\sigma_j \tag{B.13}$$

where the m_i are arbitrary integers in $\{1, \cdots, i\}$ and $P_{n_{i_1} n_{i_2}}$ are arbitrary elements of P. Then $\eta \in \Lambda_j$.

Proof. Since we have shown in Lemma B.3 that $P_{n_{i_1} n_{i_2}}(\sigma_{m_{i_1}}, \sigma_{m_{i_2}}, t)$ is a separable kernel, the kernel of the integral (B.13) is also a separable kernel. Hence, $\eta \in \Lambda_j$. □

Lemma B.4 implies that if $\hat{\lambda}_j(t|t)$ can be computed with a finite dimensional estimator for all $\lambda_j \in \Lambda_j$, then $\hat{\eta}(t|t)$ where η is defined by (B.13)) is also "*finite dimensionally computable*" (FDC).

B.2. Proofs of Theorems 2.1 and 2.2. The proofs of these two theorems are almost identical. We will prove Theorem 2.1; then we will explain how this proof is modified to prove Theorem 2.2.

Proof of Theorem 2.1. As stated in § 2, we consider the jth order Volterra term

$$\eta(t) = \int_0^t \int_0^{\sigma_1} \int_0^{\sigma_{j-1}} \gamma_1(\sigma_1) \cdots \gamma_j(\sigma_j) \cdot \xi_{k_1}(\sigma_1) \cdots \xi_{k_j}(\sigma_j)\, d\sigma_1 \cdots d\sigma_j. \tag{B.14}$$

The theorem is proved by induction on j, the order of the Volterra term. The proof for $j = 1$ is presented in § 2. We now assume the theorem holds for $j \leqq i-1$ (i.e., we assume that $E^t[e^{\xi_l(t)}\eta(t)]$ is FDC, where $\eta \in \Lambda_j$, for $j \leqq i-1$), and prove that it holds for $j = i$.

The proof is in two steps. We first reduce the problem to the computation of the elements of $\hat{\Lambda}_i$ (see Definition B.1). We then show by induction that all of the processes in $\hat{\Lambda}_i$ can be computed with finite dimensional estimators.

(i) We first consider the computation of $\hat{x}(t|t)$, where

$$x(t) = e^{\xi_l(t)}\eta(t). \tag{B.15}$$

Now

$$\hat{x}(t|t) = \int_0^t \int_0^{\sigma_1} \int_0^{\sigma_{i-1}} \gamma_1(\sigma_1) \cdots \gamma_i(\sigma_i) \cdot E^t[e^{\xi_l(t)} \xi_{k_1}(\sigma_1) \cdots \xi_{k_i}(\sigma_i)]\, d\sigma_1 \cdots d\sigma_i. \tag{B.16}$$

By (B.1) and the definition of the characteristic function, it follows that

$$\begin{aligned}
&E^t[e^{\xi_l(t)}\xi_{k_1}(\sigma_1)\cdots\xi_{k_i}(\sigma_i)] \\
&\quad = e^{\hat{\xi}_l(t|t)+(1/2)P_{ll}(t)}\{\delta_1(\sigma_1)\cdots\delta_i(\sigma_i) \\
&\qquad + \sum P_{j_1j_2}(\sigma_{m_1},\sigma_{m_2},t)\delta_{j_3}(\sigma_{m_3})\cdots\delta_{j_i}(\sigma_{m_i})+\cdots\}
\end{aligned} \tag{B.17}$$

where

$$\delta_{j_\alpha}(\sigma_{m_\alpha}) = \hat{\xi}_{j_\alpha}(\sigma_{m_\alpha}|t) + P_{l,j_\alpha}(t,\sigma_{m_\alpha},t) \tag{B.18}$$

and $\{j_\alpha, \alpha=1,\cdots,i\}$ is a permutation of $\{k_\alpha, \alpha=1,\cdots,i\}$.

Equation (B.3) implies that (B.17) can be rewritten as

$$\begin{aligned}
&E^t[e^{\xi_l(t)}\xi_{k_1}(\sigma_1)\cdots\xi_{k_i}(\sigma_i)] = e^{\hat{\xi}_l(t|t)+(1/2)P_{ll}(t)} \\
&\quad \cdot\{E^t[\xi_{k_1}(\sigma_1)\cdots\xi_{k_i}(\sigma_i)] + \sum P_{l,j_1}(t,\sigma_{m_1},t)E^t[\xi_{j_2}(\sigma_{m_2})\cdots\xi_{j_i}(\sigma_{m_i})] \\
&\qquad + \sum P_{l,j_1}(t,\sigma_{m_1},t)P_{l,j_2}(t,\sigma_{m_2},t)E^t[\xi_{j_3}(\sigma_{m_3})\cdots\xi_{j_i}(\sigma_{m_i})] \\
&\qquad + \cdots + \sum P_{l,k_1}(t,\sigma_1,t)\cdots P_{l,k_i}(t,\sigma_i,t)\}.
\end{aligned} \tag{B.19}$$

Hence, Lemmas B.2 and B.4 imply that the computation of $\hat{x}(t|t)$ involves only the computation of elements in $\hat{\Lambda}_j$, $j=1,\cdots,i$. However, the induction hypothesis implies that the elements of $\hat{\Lambda}_j$, $j=1,\cdots,i-1$ are FDC, so we need only prove that the elements of $\hat{\Lambda}_i$ are FDC.

(ii) Assume that $\eta\in\Lambda_i$ is defined by (B.14) (where $j=i$). Then the nonlinear filtering equation (A.5)–(A.6) for $\hat{\eta}(t|t)$ is

$$d\hat{\eta}(t|t) = E^t[\gamma_1(t)\xi_{k_1}(t)\lambda(t)] + \{E^t[\eta(t)\xi'(t)] - \hat{\eta}(t|t)\hat{\xi}'(t|t)\}H'(t)R^{-1}(t)\,d\nu(t) \tag{B.20}$$

where

$$d\nu(t) = dz(t) - H(t)\hat{\xi}(t|t)\,dt \tag{B.21}$$

and

$$\lambda(t) = \int_0^t\int_0^{\sigma_2}\cdots\int_0^{\sigma_{i-1}} \gamma_2(\sigma_2)\cdots\gamma_i(\sigma_i)\xi_{k_2}(\sigma_2)\cdots\xi_{k_i}(\sigma_i)\,d\sigma_2\cdots d\sigma_i \tag{B.22}$$

is an element of Λ_{i-1}; thus, by the induction hypothesis $\hat{\lambda}(t|t)$ is FDC. The first term in (B.20) (the *drift* term) is (see (B.3a))

$$\begin{aligned}
&E^t[\gamma_1(t)\xi_{k_1}(t)\lambda(t)] = \gamma_1(t)\hat{\xi}_{k_1}(t|t)\hat{\lambda}(t|t) \\
&\quad + \gamma_1(t)E^t\Big[\sum_{l=2}^i\int_0^t\int_0^{\sigma_2}\cdots\int_0^{\sigma_{i-1}} P_{k_1,k_l}(t,\sigma_i,t)\gamma_2(\sigma_2)\cdots\gamma_i(\sigma_i) \\
&\qquad \cdot\xi_{k_2}\cdots\xi_{k_{l-1}}\xi_{k_{l+1}}\cdots\xi_{k_i}\,d\sigma_2\cdots d\sigma_i\Big].
\end{aligned} \tag{B.23}$$

The first term in (B.23) is FDC by the induction hypothesis, and the second term, by Lemmas B.2 and B.4, is also FDC (i.e., it is an element of $\hat{\Lambda}_{i-2}$).

Equation (B.3a) implies that the *gain* term in (B.20) is the row vector (here $P_i(\sigma,t,t)$ denotes the ith row of $P(\sigma,t,t)$)

$$\begin{aligned}
&E^t[\eta(t)\xi'(t)] - \hat{\eta}(t|t)\hat{\xi}'(t|t) \\
&\quad = \sum_{l=1}^i E^t\Big[\int_0^t\int_o^{\sigma_1}\cdots\int_0^{\sigma_{i-1}} \gamma_1(\sigma_1)\cdots\gamma_i(\sigma_i) \\
&\qquad \cdot\xi_{k_1}(\sigma_1)\cdots\xi_{k_{l-1}}(\sigma_{l-1})\xi_{k_{l+1}}(\sigma_{l+1})\cdots\xi_{k_i}(\sigma_i)P_{k_l}(\sigma_l,t,t)\,d\sigma_1\cdots d\sigma_k\Big]
\end{aligned} \tag{B.24}$$

each element of which, by Lemmas B.2 and B.4, is an element of $\hat{\Lambda}_{i-1}$. Thus, by the induction hypothesis, the gain term, and hence the nonlinear equation (B.20) for $\hat{\eta}(t|t)$ is FDC. This completes the proof of Theorem 2.1. □

Proof of Theorem 2.2. This proof is identical to the proof of Theorem 2.1, except for the computation of the drift term in (B.20), so we will consider only that aspect of the proof. Assume that η is defined as in (2.4)—i.e., η is given by

$$\text{(B.25)} \quad \eta(t)=\int_0^t\int_0^{\sigma_1}\cdots\int_0^{\sigma_{j-1}} \xi_{k_1}(\sigma_{m_1})\cdots\xi_{k_i}(\sigma_{m_i})\gamma_1(\sigma_1)\cdots\gamma_j(\sigma_j)\,d\sigma_1\cdots d\sigma_j$$

where $i>j$; we also assume that $m_1=\cdots=m_\alpha=1$ and $m_\beta\neq 1$ for $\beta>\alpha$. In this proof, the induction is on j, the number of integrals in (B.25). That is, we assume that the theorem is true when η contains $\leqq j-1$ integrals, and prove that the theorem holds if η contains j integrals.

The nonlinear filtering equation yields

$$\text{(B.26)} \quad \begin{aligned} d\hat{\eta}(t|t)=&E^t[\gamma_1(\sigma_1)\xi_{k_1}(t)\cdots\xi_{k_\alpha}(t)\lambda(t)]\\ &+\{E^t[\eta(t)\xi'(t)]-\hat{\eta}(t|t)\hat{\xi}'(t|t)\}H'(t)R^{-1}(t)\,d\nu(t)\end{aligned}$$

where $d\nu$ is defined in (B.21) and

$$\text{(B.27)}\ \lambda(t)=\int_0^t\int_0^{\sigma_2}\cdots\int_0^{\sigma_{j-1}} \gamma_2(\sigma_2)\cdots\gamma_j(\sigma_j)\xi_{k_{\alpha+1}}(\sigma_{m_{\alpha+1}})\cdots\xi_{k_i}(\sigma_{m_i})\,d\sigma_2\cdots d\sigma_j.$$

The drift term in (B.26) is, from (B.3b),

$$\text{(B.28)} \quad \begin{aligned} &E^t[\gamma_1(t)\xi_{k_1}(t)\cdots\xi_{k_\alpha}(t)\lambda(t)]\\ &\quad=\gamma_1(t)E^t[\xi_{k_1}(t)\cdots\xi_{k_\alpha}(t)]\hat{\lambda}(t|t)\\ &\qquad+\gamma_1(t)\sum\Big\{E^t[\xi_{l_2}(t)\cdots\xi_{l_\alpha}(t)]\\ &\qquad\quad\cdot E^t\Big[\int_0^t\int_0^{\sigma_2}\cdots\int_0^{\sigma_{j-1}}\gamma_2(\sigma_2)\cdots\gamma_j(\sigma_j)P_{l_1l_{\alpha+1}}(t,\sigma_{m_{\alpha+1}},t)\\ &\qquad\qquad\cdot\xi_{l_{\alpha+2}}(\sigma_{m_{\alpha+2}})\cdots\xi_{l_i}(\sigma_{m_i})\,d\sigma_2\cdots d\sigma_j\Big]\Big\}+\cdots\end{aligned}$$

where $\{l_1,\cdots,l_\alpha\}$ is a permutation of $\{k_1,\cdots,k_\alpha\}$ and $\{l_{\alpha+1},\cdots,l_i\}$ is a permutation of $\{k_{\alpha+1},\cdots,k_i\}$. The first term of (B.28) is FDC by the induction hypothesis, and the other terms, by Lemmas B.2 and B.4 and the induction hypothesis, are also FDC. We have also used the fact that the conditional distribution of $\xi(t)$ given z^t is Gaussian (Lemma 2.1) in order to conclude that $E^t[\xi_{k_1}(t)\cdots\xi_{k_\alpha}(t)]$ can be computed (via (B.3c)) as a memoryless function of $\hat{\xi}(t|t)$ and $P(t)$.

The gain term in (B.26) is also FDC; the proof is identical to that of Theorem 2.1. Hence $\hat{\eta}(t|t)$ is FDC, and Theorem 2.2 is proved. □

Acknowledgment. The authors would like to thank Professor Roger Brockett of Harvard University for many helpful discussions and for suggesting the use of Volterra series in the present context.

REFERENCES

[1] R. W. BROCKETT, *System theory on group manifolds and coset spaces*, SIAM J. Control, 10 (1972), pp. 265–284.

[2] ———, *Volterra series and geometric control theory*, Automatica, 12 (1976), pp. 167–176.

[3] ———, *Nonlinear systems and differential geometry*, Proc. IEEE, 64 (January 1976), pp. 61–72.
[4] ———, Personal communication.
[5] P. D'ALESSANDRO, A. ISIDORI AND A. RUBERTI, *Realizations and structure theory of bilinear dynamical systems*, SIAM J. Control, 12 (1974), pp. 517–535.
[6] M. FLIESS, *Un outil algebrique: les series formelles non commutatives*, Mathematical Systems Theory, G. Marchesini and S. K. Mitter, eds., Springer-Verlag, New York, 1976, pp. 122–148.
[7] M. FUJISAKI, G. KALLIANPUR, AND H. KUNITA, *Stochastic differential equations for the nonlinear filtering problem*, Osaka J. Math., 9 (1972), pp. 19–40.
[8] A. H. JAZWINSKI, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.
[9] C. JOHNSON AND E. B. STEAR, *Optimal filtering in the presence of multiplicative noise*, Proc. Fifth Symp. Nonlinear Estimation Theory and Its Applications, September 1974, Western Periodicals, North Hollywood, CA, pp. 124–134.
[10] V. JURDJEVIC AND H. J. SUSSMAN, *Control systems on Lie groups*, J. Differential Equations, 12 (1972), pp. 313–329.
[11] T. KAILATH AND P. FROST, *An innovations approach to least-squares estimation, Part II: Linear smoothing in additive white noise*, IEEE Trans. Automatic Control, AC-13 (1968), pp. 655–660.
[12] R. E. KALMAN AND R. S. BUCY, *New results in linear filtering and prediction theory*, J. Basic Engr. (Trans. ASME), 83D (1961) pp. 95–108.
[13] A. J. KRENER, *On the equivalence of control systems and linearization of nonlinear systems*, SIAM J. Control, 11 (1973), pp. 670–676.
[14] H. J. KUSHNER, *Dynamical equations for optimal nonlinear filtering*, J. Differential Equations, 3 (1967), pp. 179–190.
[15] H. KWAKERNAAK, *Optical filtering in linear systems with time delays*, IEEE Trans. Automatic Control, AC-12 (1967), pp. 169–173.
[16] C. T. LEONDES, J. B. PELLER AND E. B. STEAR, *Nonlinear smoothing theory*, IEEE Trans. Systems Sci. Cybernet., SSC-6 (1970), p. 63–71.
[17] J. T. LO AND A. S. WILLSKY, *Estimation for rotational processes with one degree of freedom I: introduction and continuous time processes*, IEEE Trans. Automatic Control, AC-20 (1975), pp. 10–21.
[18] S. I. MARCUS, *Estimation and analysis of nonlinear stochastic systems*, Ph.D. thesis, Dept. of Electrical Engineering, Mass. Inst. of Tech., Cambridge, June 1975; M.I.T. Electronic Systems Laboratory Rep. ESL-R-601, June 1975.
[19] S. I. MARCUS AND A. S. WILLSKY, *A class of finite dimensional optimal nonlinear estimators*, Proc. Fifth Symp. Nonlinear Estimation Theory and Its Applications, September 1974, Western Periodicals, North Hollywood, CA, pp. 158–168.
[20] D. N. MARTIN, *Stability criteria for systems with colored multiplicative noise*, Ph.D. thesis, Dept. of Electrical Engineering, Mass. Inst. of Tech. Cambridge, June 1974; M.I.T. Electronic Systems Laboratory Rep. ESL-R-567, June 1974.
[21] K. S. MILLER, *Multidimensional Gaussian Distributions*, John Wiley, New York, 1965.
[22] I. B. RHODES, *A tutorial introduction to estimation and filtering*, IEEE Trans. Automatic Control, AC-16 (1971), pp. 688–706.
[23] A. A. SAGLE AND R. E. WALDE, *Introduction to Lie Groups and Lie Algebras*, Academic Press, New York, 1973.
[24] H. J. SUSSMANN, *Semigroup representations, bilinear approximation of input-output maps, and generalized inputs*, Mathematical Systems Theory, G. Marchesini and S. K. Mitter, eds., Springer-Verlag, New York, 1976, pp. 172–192.
[25] A. S. WILLSKY, *Some estimation problems on Lie groups*, Geometric Methods in System Theory, R. W. Brockett and D. W. Mayne, eds., Reidel, Hingham, MA, 1973.
[26] A. S. WILLSKY AND S. I. MARCUS, *Estimation for bilinear stochastic systems*, Proc. of the 1974 Conf. on Bilinear Systems, Lecture Notes in Economics and Mathematical Systems, R. R. Mohler and A. Ruberti, eds., Springer-Verlag, New York, 1975, pp. 116–137.
[27] ———, *Analysis of bilinear noise models in circuits and devices*, J. Franklin Inst., 301 (1976), pp. 103–122.
[28] W. M. WONHAM, *On a matrix Riccati equation of stochastic control*, SIAM J. Control, 6 (1968), pp. 681–697.

SIAM J. MATH. ANAL.
Vol. 9, No. 2, April, 1978
Copyright © 1978 Society for Industrial and Applied Mathematics

SUFFICIENT CONDITIONS FOR ACKERBERG–O'MALLEY RESONANCE*

F. W. J. OLVER†

Abstract. An investigation is made of the asymptotic nature of the solution of the boundary-value problem

$$\varepsilon y''+2xA(\varepsilon, x)y'-A(\varepsilon, x)B(\varepsilon, x)y=0; \qquad y(a)=l, \qquad y(b)=m,$$

as $\varepsilon \to 0$, where $A(\varepsilon, x)$ and $B(\varepsilon, x)$ are continuous real functions of ε and x, $a<0$, $b>0$, and $A(\varepsilon, x)$ is nonzero in $[a, b]$. Particular attention is paid to the problem of resonance, which arises when the limiting form of the solution exhibits an unusual lack of decay (in the case $A(\varepsilon, x)<0$), or an unusual rate of growth (in the case $A(\varepsilon, x)>0$). By application of a recent theory of differential equations with coalescing turning points sufficient conditions for resonance are established, both with and without the assumption that $A(\varepsilon, x)$ and $B(\varepsilon, x)$ are analytic functions of ε and x. Illustrative examples are also included.

1. Introduction. In this paper we study the differential equation

$$\varepsilon y''+2xA(\varepsilon, x)y'-A(\varepsilon, x)B(\varepsilon, x)y=0, \tag{1.01}$$

in which the independent variable x ranges over a compact interval $[a, b]$ with $a<0$ and $b>0$, and ε is a small positive parameter. The coefficients $A(\varepsilon, x)$ and $B(\varepsilon, x)$ are continuous real functions of ε and x, and $A(\varepsilon, x)$ is nonzero in $[a, b]$. The boundary conditions are assumed to be

$$y(a)=l, \qquad y(b)=m, \tag{1.02}$$

where l and m are prescribed real constants, at least one of which is nonzero. The problem stems from the flow of viscous fluid between rotating disks [17], [5], and interest is focused on the asymptotic behavior of the solution y as $\varepsilon \to 0$.

By assuming that the coefficients $A(\varepsilon, x)$ and $B(\varepsilon, x)$ are independent of ε, or, more generally, expandable in asymptotic power series of the form

$$\begin{aligned} A(\varepsilon, x) &\sim A_0(x)+A_1(x)\varepsilon+A_2(x)\varepsilon^2+\cdots, \\ B(\varepsilon, x) &\sim B_0(x)+B_1(x)\varepsilon+B_2(x)\varepsilon^2+\cdots, \end{aligned} \tag{1.03}$$

the problem can be treated by the methods of singular perturbation theory [15]. The nature of the solution depends on the sign of $A(\varepsilon, x)$ in $[a, b]$. When this sign is negative (which is the case of greater physical interest), it is found that in general $y \to 0$ as $\varepsilon \to 0$, uniformly in any closed interval that is properly interior to $[a, b]$. However, certain exceptional situations may arise in which the solution does not decay in this manner. Ackerberg and O'Malley drew attention to this phenomenon in [1] and called it *resonance*. Using two distinct methods, these writers (and also O'Malley [13] and [14, Chap. 8]), showed that a necessary condition for resonance to take place is given by

$$B(0, 0)=2s-2, \qquad s=1, 2, 3, \cdots, \tag{1.04}$$

and they included an example in which resonance actually occurs. The first method of Ackerberg and O'Malley is to construct the Liouville–Green (or WKBJ) approximations to the solutions of (1.01) valid in intervals that exclude the turning point at $x=0$ and to link these solutions by skirting the turning point in the complex plane by the method of Zwaan [19]. The second method is based on the uniform reduction theorems

* Received by the editors June 1, 1976, and in revised form October 12, 1976.

† Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, and National Bureau of Standards, Washington, DC 20234. This research was supported by the U.S. Army Research Office, Durham, under Contract DA ARO D 31 124 73 G204, and by the National Science Foundation under Grant GP 32841X2.

of Sibuya [16] and Lee [6], and transforms (1.01) into a differential equation of the same type in which $A(\varepsilon, x)$ and $B(\varepsilon, x)$ are replaced by functions that are independent of x. This yields approximations for the solutions in terms of parabolic cylinder functions that are uniformly valid throughout $[a, b]$. The first method was applied in a formal manner, but the second method was placed on a rigorous foundation by assuming that the coefficients $A(\varepsilon, x)$ and $B(\varepsilon, x)$ are analytic functions of ε and x.

The first method of Ackerberg and O'Malley was reconsidered independently by Watts [17] and Zauderer [18]. Avoiding Zwaan's complex-variable approach, these writers showed that the linking of the Liouville–Green approximations across the turning point can be achieved with the aid of local approximations in terms of parabolic cylinder functions. These local approximations are obtained by making a stretching transformation of the independent variable.

Next, Lakin [5] and Cook and Eckhaus [2] demonstrated independently, by formal methods, that the condition (1.04) is only the first of an infinite sequence of necessary conditions for resonance. Lakin constructed series solutions in terms of parabolic cylinder functions that are uniformly valid in $[a, b]$, whereas Cook and Eckhaus used the method of matched asymptotic expansions.

More recently, Kreiss and Parter [4] and Matkowsky [7], [8] have considered the problem and discussed several examples in detail. The approach of Kreiss and Parter is quite different from the approaches of other writers that have been mentioned, and is based in part on the maximum principle. Matkowsky uses a formal criterion that does not depend on Zwaan's method or on local approximations in the neighborhood of $x=0$. We shall discuss some of Matkowsky's results more fully later (§ 9).

A similar situation exists when the coefficient $A(\varepsilon, x)$ in (1.01) is positive throughout $[a, b]$. In [13], and also [14, Chap. 8], O'Malley showed that in general the solution y is bounded as $\varepsilon \to 0$, except in the neighborhood of $x=0$ where it may grow algebraically, that is, at the rate of a negative power of ε. However, in certain exceptional cases the solution grows exponentially in closed intervals that are properly interior to $[a, 0]$ or $[0, b]$. These exceptional cases are again regarded as manifestations of a resonance phenomenon, and O'Malley proved that when the coefficients $A(\varepsilon, x)$ and $B(\varepsilon, x)$ are analytic in ε and x a necessary condition for resonance is given by

$$(1.05) \qquad B(0,0)=-2s, \qquad s=1,2,3,\cdots.$$

Finally, de Groen [3] has considered resonance problems of the same general type for certain elliptic partial differential equations.

None of the references cited establishes sufficient conditions for resonance that are directly applicable to the general case. Furthermore, the only ordinary differential equations for which resonance has been conclusively demonstrated are either solvable exactly in terms of parabolic cylinder functions, or have a nontrivial solution that is independent of ε. In § 8 of the present paper we shall remedy this situation by establishing a sufficient condition for resonance, applicable to each of the cases $A(\varepsilon, x)<0$ and $A(\varepsilon, x)>0$. In effect this condition requires (1.01) to be transformable into a particular form of Weber's equation (that is, the differential equation satisfied by the parabolic cylinder functions), except for the inclusion of a coefficient term that may depend x as well as ε and is uniformly and exponentially small in $[a, b]$ as $\varepsilon \to 0$. Not surprisingly, the particular form of Weber's equation has to be one that admits a solution that is recessive both for large positive and large negative arguments. We also give new examples in which the sufficiency condition is satisfied.

The main purpose of the present paper, however, is to reconsider the problem without the requirement that $A(\varepsilon, x)$ and $B(\varepsilon, x)$ be analytic functions of ε and x. By permitting nonanalytic behavior, we shall see that resonance occurs without the need for (1.01) to be transformable into the special form of differential equation that differs from Weber's equation only by an exponentially small term in the coefficient. In fact, from the new standpoint resonance occurs in the general case just as frequently as it does with Weber's equation. This is a more natural state of affairs because with the assumptions (1.03) the original differential equation (1.01) is always transformable into an equation having the same asymptotic structure as Weber's equation.

Another way of viewing the situation is that by imposing analytic behavior on the coefficients $A(\varepsilon, x)$ and $B(\varepsilon, x)$, the boundary-value problem becomes ill-posed when the condition (1.04) (or (1.05)) is satisfied. That is, nonresonant solutions are transformable into resonant solutions by making relatively small numerical changes in the coefficients.

Our approach will be to introduce an extra free parameter κ in the differential equation, which we define by

$$\pm\kappa = 1 + B(0, 0), \tag{1.06}$$

the upper sign applying in the case in which $A(\varepsilon, x)$ is negative, and the lower sign when $A(\varepsilon, x)$ is positive. In order to facilitate applications of the uniform asymptotic theory of linear differential equations, the term in the first derivative in (1.01) is removed by change of dependent variable. From the standpoint of this theory the transformed equation may be regarded as having two simple turning points which depend on ε and κ and coalesce into a double turning point as $\varepsilon \to 0$. This is exactly the situation treated in a recent paper by the present writer [12], and application of the theory given in this reference yields asymptotic solutions of (1.01) for small ε that are uniformly valid with respect to $x \in [0, b]$ or $x \in [a, 0]$, and also with respect to positive values of the parameter κ that are bounded and bounded away from zero. The approximate solutions are expressed in terms of parabolic cylinder functions, accompanied by strict error bounds. On combining the solutions in such a way that the boundary conditions (1.02) are satisfied, we find that the condition for resonance becomes a type of eigenvalue problem resulting in a transcendental equation for κ. The solutions $\kappa = \kappa(\varepsilon)$ of this equation form a continuum, and are therefore not analytic functions of ε. However, $\kappa(\varepsilon)$ can be approximated asymptotically, and the results are consistent with the necessary conditions for resonance given by (1.04) when $A(\varepsilon, x)$ is negative, or (1.05) when $A(\varepsilon, x)$ is positive.

2. Assumptions and preliminary transformations. In the given differential equation (1.01) we shall assume that

$$A(\varepsilon, x) = A_0(x) + A_1(x)\varepsilon + A_2(\varepsilon, x)\varepsilon^2, \tag{2.01}$$

$$B(\varepsilon, x) = \pm\kappa - 1 + x\hat{B}_0(x) + B_1(\varepsilon, x)\varepsilon, \tag{2.02}$$

subject to the following conditions:

(i) $A_0(x)$, $A_1(x)$, and $\hat{B}_0(x)$ are independent of ε.

(ii) $A_0(x)$ is nonvanishing in $[a, b]$, and the upper or lower sign is taken in (2.02) according as $A_0(x)$ is negative or positive.

(iii) $A_0^{\mathrm{vi}}(x)$, $A_1^{\mathrm{v}}(x)$, and $\hat{B}_0^{\mathrm{v}}(x)$ are continuous in $[a, b]$.[1]

(iv) $A_2(\varepsilon, x)$, $\partial A_2(\varepsilon, x)/\partial x$, and $B_1(\varepsilon, x)$ are continuous functions of ε and x when $\varepsilon \in [0, \delta]$ and $x \in [a, b]$, where δ is a positive constant.

[1] As usual, $A_0^{\mathrm{vi}}(x)$ denotes the sixth derivative of $A_0(x)$, and so on. These conditions are stronger than actually needed, but they simplify the exposition.

(v) κ is a free parameter with the range $k_1 \leqq \kappa \leqq k_2$, where k_1 and k_2 are constants such that $0 < k_1 < k_2 < \infty$.

We remove the term in the first derivative from the differential equation (1.01) by taking a new dependent variable $w = w(\varepsilon, \kappa, x)$, defined by

$$w = \exp\left\{\frac{1}{\varepsilon}\int_0^x tA(\varepsilon, t)\,dt\right\}y. \tag{2.03}$$

Then

$$w'' = \varepsilon^{-2}\{x^2A^2 + \varepsilon(xA)' + \varepsilon AB\}w, \tag{2.04}$$

where primes again denote partial differentiations with respect to x. Substituting for A and B by means of (2.01) and (2.02) and rearranging, we find that

$$w'' = \{\varepsilon^{-2}f(\varepsilon, \kappa, x) + g(\varepsilon, \kappa, x)\}w, \tag{2.05}$$

where

$$f(\varepsilon, \kappa, x) = x^2A_0^2 + \{xA_0' \pm \kappa A_0 + 2x^2A_0A_1 + xA_0\hat{B}_0\}\varepsilon, \tag{2.06}$$

$$\begin{aligned} g(\varepsilon, \kappa, x) = {} & x^2A_1^2 + xA_1' \pm \kappa A_1 + 2x^2A_0A_2 + xA_1\hat{B}_0 + A_0B_1 \\ & + (2x^2A_1A_2 + xA_2' \pm \kappa A_2 + xA_2\hat{B}_0 + A_1B_1)\varepsilon \\ & + (x^2A_2^2 + A_2B_1)\varepsilon^2. \end{aligned} \tag{2.07}$$

In consequence of conditions (iii) and (iv), $f(\varepsilon, \kappa, x)$, $g(\varepsilon, \kappa, x)$, and the first five partial x-derivatives of $f(\varepsilon, \kappa, x)$ are continuous functions of ε, κ, and x.

To find the boundary conditions satisfied by w, we introduce the notations

$$P(x) = \int_0^x tA_0(t)\,dt, \qquad Q(x) = \int_0^x tA_1(t)\,dt. \tag{2.08}$$

Then from (2.01) we have

$$\frac{1}{\varepsilon}\int_0^x tA(\varepsilon, t)\,dt = \frac{1}{\varepsilon}P(x) + Q(x) + O(\varepsilon), \tag{2.09}$$

as $\varepsilon \to 0$, uniformly with respect to $x \in [a, b]$. Hence from (1.02) and (2.03) the new boundary conditions are given by

$$w(\varepsilon, \kappa, a) = le^{Q(a)}\,e^{P(a)/\varepsilon}\{1 + O(\varepsilon)\}, \tag{2.10}$$

$$w(\varepsilon, \kappa, b) = me^{Q(b)}\,e^{P(b)/\varepsilon}\{1 + O(\varepsilon)\}. \tag{2.11}$$

3. Zeros of $f(\varepsilon, \kappa, x)$. Write

$$\phi(x) = |A_0(x)|, \qquad \omega(x) = \{A_0'(x) + 2xA_0(x)A_1(x) + A_0(x)\hat{B}_0(x)\}/\phi(x). \tag{3.01}$$

Then in $[a, b]$, $\phi(x)$ is nonvanishing and six times continuously differentiable, and $\omega(x)$ is five times continuously differentiable. From (2.06) we see that

$$f(\varepsilon, \kappa, x) = x^2\phi^2(x) - \{\kappa - x\omega(x)\}\phi(x)\varepsilon. \tag{3.02}$$

Let r and R denote the positive constants

$$r = \min_{a \leqq x \leqq b} \phi(x), \qquad R = \max_{a \leqq x \leqq b} |\omega(x)|.$$

In the interval $-k_1/(2R) \leqq x \leqq k_1/(2R)$ the factor $\kappa - x\omega(x)$ is positive; accordingly the zeros of $f(\varepsilon, \kappa, x)$ are given by

$$\pm\Xi(\kappa, x)=\varepsilon^{1/2},$$

where here and elsewhere $\varepsilon^{1/2}$ denotes as usual the positive square root, and

$$\Xi(\kappa, x)=x\{\phi(x)\}^{1/2}\{\kappa-x\omega(x)\}^{-1/2}.$$

It is easily seen that $\Xi(\kappa, x)$ is continuous and $\partial\Xi(\kappa, x)/\partial x$ is positive and continuous in the region $k_1\leqq\kappa\leqq k_2$, $-k\leqq x\leqq k$, where k is an assignable positive constant. We now apply the implicit function theorem in the form given in [12, p. 144], with κ in the role of a and $\varepsilon^{1/2}$ in the role of ξ. We deduce that as long as the constant δ is small enough, the only zeros of $f(\varepsilon, \kappa, x)$ in the interval $-k\leqq x\leqq k$ are given by $x=z_1(\varepsilon, \kappa)$ and $x=z_2(\varepsilon, \kappa)$, where $z_1(\varepsilon, \kappa)$ and $z_2(\varepsilon, \kappa)$ are continuous functions of ε and κ, $z_1(\varepsilon, \kappa)$ being nonpositive and a decreasing function of ε, and $z_2(\varepsilon, \kappa)$ being nonnegative and an increasing function of ε; furthermore

$$z_1(0, \kappa)=z_2(0, \kappa)=0. \tag{3.03}$$

Again, by taking δ small enough we can also ensure that $z_1(\varepsilon, \kappa)$ and $z_2(\varepsilon, \kappa)$ are the only zeros of $f(\varepsilon, \kappa, x)$ in the whole of the original interval $[a, b]$; in fact it suffices that

$$\delta<rk^2/\{k_2+R\max(|a|, b)\}.$$

The asymptotic forms of the zeros $z_1(\varepsilon, \kappa)$ and $z_2(\varepsilon, \kappa)$ are found by equating the right-hand side of (3.02) to zero, and solving by standard asymptotic methods given, for example, in [11, p. 13]. We find that

$$\begin{aligned} z_1(\varepsilon, \kappa)&=-\left\{\frac{\kappa\varepsilon}{\phi(0)}\right\}^{1/2}-\left\{\omega(0)+\kappa\frac{\phi'(0)}{\phi(0)}\right\}\frac{\varepsilon}{2\phi(0)}+O(\varepsilon^{3/2}),\\ z_2(\varepsilon, \kappa)&=\left\{\frac{\kappa\varepsilon}{\phi(0)}\right\}^{1/2}-\left\{\omega(0)+\kappa\frac{\phi'(0)}{\phi(0)}\right\}\frac{\varepsilon}{2\phi(0)}+O(\varepsilon^{3/2}), \end{aligned} \tag{3.04}$$

as $\varepsilon\to 0$, uniformly with respect to κ.

We shall need certain properties of the function $p(\varepsilon, \kappa, x)$, defined by

$$p(\varepsilon, \kappa, x)=f(\varepsilon, \kappa, x)/\{(x-z_1)(x-z_2)\}, \tag{3.05}$$

when $x\neq z_1$ or z_2, or by the limiting form of this quotient when $x=z_1$ or z_2. Clearly $p(\varepsilon, \kappa, x)$ is always positive, and from (3.02) and (3.03) we have

$$p(0, \kappa, x)=\phi^2(x). \tag{3.06}$$

LEMMA 3.1. *The function $p(\varepsilon, \kappa, x)$ and its first three partial x-derivatives are continuous in the cuboid*

$$0\leqq\varepsilon\leqq\delta, \qquad k_1\leqq\kappa\leqq k_2, \qquad a\leqq x\leqq b. \tag{3.07}$$

In the proof that follows we suppress the arguments ε and κ of f and p in most places; thus $f(x)\equiv f(\varepsilon, \kappa, x)$ and $p(x)\equiv p(\varepsilon, \kappa, x)$. We continue to use primes to denote partial differentiations with respect to x.

Since $f(x)$ is a continuous function of ε, κ, and x, and z_1 and z_2 are continuous functions of ε and κ, it follows immediately that $p(x)$ is continuous throughout the cuboid (3.07), except possibly at those points for which $x=z_1$ or z_2. Let us define

$$h(x)=f(x)/(x-z_1), \qquad x\neq z_1; \qquad h(z_1)=f'(z_1). \tag{3.08}$$

Then

$$p(x)=h(x)/(x-z_2), \qquad x\neq z_1 \text{ or } z_2. \tag{3.09}$$

By the mean-value theorem $h(x)=f'(c_1)$, where c_1 lies between x and z_1. Since $f'(x)$ is continuous (in all variables) it follows that $h(x)$ is continuous when $x=z_1$, and therefore throughout (3.07). If $\varepsilon\neq 0$, then $z_1\neq z_2$ and it follows from (3.09) that $p(x)$ is continuous when $x=z_1$, and therefore, by symmetry, when $x=z_2$.

It remains to consider the edge $\varepsilon=x=0$ in order to establish the continuity of $p(x)$ throughout the whole of (3.07). Differentiating (3.08), we have

$$h'(x)=\{(x-z_1)f'(x)-f(x)\}/(x-z_1)^2, \qquad x\neq z_1. \tag{3.10}$$

From Taylor's theorem

$$f'(x)=f'(z_1)+(x-z_1)f''(c_2), \qquad f(x)=(x-z_1)f'(z_1)+\tfrac{1}{2}(x-z_1)^2f''(c_3),$$

where c_2 and c_3 lie between x and z_1. Substitution in (3.10) yields

$$h'(x)=f''(c_2)-\tfrac{1}{2}f''(c_3), \qquad x\neq z_1. \tag{3.11}$$

From (3.08) we have by definition

$$h'(z_1)=\lim_{x\to z_1}\left\{\frac{f(x)}{x-z_1}-f'(z_1)\right\}\frac{1}{x-z_1}=\tfrac{1}{2}f''(z_1). \tag{3.12}$$

The last two results show that $h'(x)$ is continuous when $x=z_1$, and therefore throughout (3.07). Now suppose that $\varepsilon\neq 0$. Since $h(z_2)=0$, application of the mean-value theorem gives

$$h(x)=(x-z_2)h'(d_1),$$

where d_1 lies between x and z_2. Substituting for $h'(d_1)$ by means of (3.11) and (3.12), we obtain

$$h(x)=(x-z_2)\{f''(e_1)-\tfrac{1}{2}f''(e_2)\},$$

where e_1 and e_2 lie in the smallest interval I, say, that contains x, z_1, and z_2. Hence

$$p(x)=f''(e_1)-\tfrac{1}{2}f''(e_2), \tag{3.13}$$

provided that $x\neq z_2$. By symmetry, however, (3.13) also holds when $x\neq z_1$. Therefore (3.13) applies at every point of (3.07) except those on the face $\varepsilon=0$. From (3.02) and (3.06) it follows that $p(x)=\frac{1}{2}f''(0)$ when $\varepsilon=x=0$. Because $f''(x)$ is continuous, it follows from (3.13) that $p(x)$ is continuous at all points of the edge $\varepsilon=x=0$. This completes the part of the proof concerning the continuity of $p(\varepsilon,\kappa,x)$.

The proofs for the derivatives $\partial p(\varepsilon,\kappa,x)/\partial x$, $\partial^2p(\varepsilon,\kappa,x)/\partial x^2$, and $\partial^3 p(\varepsilon,\kappa,x)/\partial x^3$ are similar and it is unnecessary to record details, except perhaps the equations that correspond to (3.13):

$$p'(x)=\tfrac{11}{6}f'''(e_3)-\tfrac{5}{3}f'''(e_4), \qquad p''(x)=\tfrac{53}{12}f^{\text{iv}}(e_5)-\tfrac{13}{3}f^{\text{iv}}(e_6),$$
$$p'''(x)=\tfrac{233}{20}f^{\text{v}}(e_7)-\tfrac{58}{5}f^{\text{v}}(e_8),$$

where $e_3, e_4, \cdots, e_8$ are points in the interval I defined above. In particular, when $\varepsilon=x=0$ we have

$$p'(0)=\tfrac{1}{6}f'''(0), \qquad p''(0)=\tfrac{1}{12}f^{\text{iv}}(0), \qquad p'''(0)=\tfrac{1}{20}f^{\text{v}}(0).$$

LEMMA 3.2. *As* $\varepsilon \to 0$

$$p(\varepsilon, \kappa, x) = \phi^2(x) + O(\varepsilon), \tag{3.14}$$

uniformly with respect to $\kappa \in [k_1, k_2]$ *and* $x \in [a, b]$.

Since $p(0, \kappa, x)$ equals $\phi^2(x)$, in effect this result supplies information concerning the partial derivative $\partial p(\varepsilon, \kappa, x)/\partial\varepsilon$. The proof is as follows. Write

$$\hat{f}(x) \equiv \hat{f}(\varepsilon, \kappa, x) = f(\varepsilon, \kappa, x) - (x - z_1)(x - z_2)\phi^2(x), \tag{3.15}$$

$$\hat{p}(x) \equiv \hat{p}(\varepsilon, \kappa, x) = \hat{f}(\varepsilon, \kappa, x)/\{(x - z_1)(x - z_2)\}. \tag{3.16}$$

Then $\hat{f}(x)$ and its first five partial x-derivatives are continuous in the cuboid (3.07) and $\hat{f}(x)$ vanishes when $x = z_1$ or z_2. Lemma 3.1 may now be applied with $f(x)$ and $p(x)$ replaced by $\hat{f}(x)$ and $\hat{p}(x)$, respectively, and *inter alia* we deduce that

$$\hat{p}(x) = \hat{f}''(\hat{e}_1) - \tfrac{1}{2}\hat{f}''(\hat{e}_2), \tag{3.17}$$

where $\hat{e}_1$ and $\hat{e}_2$ lie in I; compare (3.13). From (3.02) and (3.15) we see that

$$\hat{f}(x) = \{(z_1 + z_2)x - z_1 z_2\}\phi^2(x) - \{\kappa - x\omega(x)\}\phi(x)\varepsilon. \tag{3.18}$$

The approximations (3.04) show that as $\varepsilon \to 0$ both $z_1 + z_2$ and $z_1 z_2$ are uniformly $O(\varepsilon)$. From this result and the twice-differentiated form of (3.18) it is evident that as $\varepsilon \to 0$, $\hat{f}''(x)$ is uniformly $O(\varepsilon)$; hence from (3.17) $\hat{p}(x)$ is uniformly $O(\varepsilon)$. Combination of the last result with (3.05), (3.15), and (3.16) yields the desired relation (3.14).

4. Application of the Liouville transformation. Returning to the differential equation (2.05), we follow [12, § 2], and take a new independent variable ζ, defined by

$$\dot{x}^2 f(\varepsilon, \kappa, x) = \zeta^2 - \alpha^2, \tag{4.01}$$

where the dot signifies differentiation with respect to ζ, and α is a nonnegative real number chosen to make $\zeta = -\alpha$ correspond to $x = z_1$, and $\zeta = \alpha$ correspond to $x = z_2$. Thus $\alpha \equiv \alpha(\varepsilon, \kappa)$ is given by

$$\int_{z_1}^{z_2} \{-f(\varepsilon, \kappa, x)\}^{1/2}\, dx = \int_{-\alpha}^{\alpha} (\alpha^2 - \zeta^2)^{1/2}\, d\zeta = \tfrac{1}{2}\pi\alpha^2. \tag{4.02}$$

Next, integration of (4.01) yields the following relations between x and ζ:

$$\int_{z_1}^{x} \{-f(\varepsilon, \kappa, t)\}^{1/2}\, dt = \int_{-\alpha}^{\zeta} (\alpha^2 - \tau^2)^{1/2}\, d\tau, \qquad z_1 \leqq x \leqq z_2, \tag{4.03}$$

$$\int_{z_2}^{x} \{f(\varepsilon, \kappa, t)\}^{1/2}\, dt = \int_{\alpha}^{\zeta} (\tau^2 - \alpha^2)^{1/2}\, d\tau, \qquad z_2 \leqq x \leqq b, \tag{4.04}$$

$$\int_{x}^{z_1} \{f(\varepsilon, \kappa, t)\}^{1/2}\, dt = \int_{\zeta}^{-\alpha} (\tau^2 - \alpha^2)^{1/2}\, d\tau, \qquad a \leqq x \leqq z_1. \tag{4.05}$$

These equations define ζ as a continuous and increasing function of x in $[a, b]$. We suppose that $\zeta = \bar{a}$ (<0) corresponds to $x = a$, and $\zeta = \bar{b}$ (>0) corresponds to $x = b$.

The Liouville transformation of (2.05) is completed by introducing a new dependent variable W, defined by

$$W = \dot{x}^{-1/2} w. \tag{4.06}$$

Then

$$d^2W/d\zeta^2 = \{\varepsilon^{-2}(\zeta^2 - \alpha^2) + \psi(\alpha, \kappa, \zeta)\}W, \tag{4.07}$$

where

$$\psi(\alpha, \kappa, \zeta) = \dot{x}^2 g(\varepsilon, \kappa, x) + \dot{x}^{1/2} d^2(\dot{x}^{-1/2})/d\zeta^2. \tag{4.08}$$

LEMMA 4.1. *Provided that δ is sufficiently small, $\alpha(\varepsilon, \kappa)$ is a continuous function of ε and κ, and an increasing function of ε in the rectangle*

$$0 \leqq \varepsilon \leqq \delta, \qquad k_1 \leqq \kappa \leqq k_2, \tag{4.09}$$

and $\psi(\alpha, \kappa, \zeta)$ is a continuous function of α, κ, and ζ throughout the region

$$0 \leqq \alpha \leqq \alpha(\delta, \kappa), \qquad k_1 \leqq \kappa \leqq k_2, \qquad \bar{a} \leqq \zeta \leqq \bar{b}. \tag{4.10}$$

For fixed values of κ we obtain this result by applying Lemma I of [12], with ε^{-1} in the role of u, $\frac{1}{2}(z_2 - z_1)$ in the role of a, and $x - \frac{1}{2}(z_1 + z_2)$ in the role of x, and using Lemma 3.1 above to verify that the requisite conditions are satisfied. The added properties of continuity with respect to ε and κ in the case of $\alpha(\varepsilon, \kappa)$, or continuity with respect of α, κ, and ζ in the case of $\psi(\alpha, \kappa, \zeta)$, are verifiable by repeating the steps in the proof of Lemma I of [12] in a straightforward manner.

We shall need asymptotic approximations to the values of α and ζ when ε is small. These estimates are supplied by the next two lemmas.

LEMMA 4.2. *As $\varepsilon \to 0$*

$$\alpha^2 = \kappa\varepsilon + O(\varepsilon^2), \tag{4.11}$$

uniformly with respect to $\kappa \in [k_1, k_2]$.

On substituting in (4.02) by means of (3.05), we obtain

$$\alpha^2 = \frac{2}{\pi} \int_{z_1}^{z_2} \{(z_2 - x)(x - z_1)\}^{1/2} \{p(\varepsilon, \kappa, x)\}^{1/2}\, dx.$$

When $x \in [z_1, z_2]$, we have $x = O(\varepsilon^{1/2})$; compare (3.04). Hence from Taylor's theorem and Lemmas 3.1 and 3.2, we find that

$$\begin{aligned} p(\varepsilon, \kappa, x) &= p(\varepsilon, \kappa, 0) + x p'(\varepsilon, \kappa, 0) + O(x^2) \\ &= \phi^2(0) + x p'(\varepsilon, \kappa, 0) + O(\varepsilon), \end{aligned} \tag{4.12}$$

the last O-term being uniform with respect to κ and x. Therefore

$$\alpha^2 = \frac{2}{\pi} \int_{z_1}^{z_2} \{(z_2 - x)(x - z_1)\}^{1/2} \left\{ \phi(0) + \frac{x p'(\varepsilon, \kappa, 0)}{2\phi(0)} + O(\varepsilon) \right\} dx. \tag{4.13}$$

We now make the substitutions

$$x = \tfrac{1}{2}(z_1 + z_2) + v, \qquad z = \tfrac{1}{2}(z_2 - z_1), \tag{4.14}$$

and note that

$$z = \{\kappa\varepsilon/\phi(0)\}^{1/2} + O(\varepsilon^{3/2}); \tag{4.15}$$

compare (3.04). Then

$$\frac{2}{\pi} \int_{z_1}^{z_2} \{(z_2 - x)(x - z_1)\}^{1/2}\, dx = \frac{2}{\pi} \int_{-z}^{z} (z^2 - v^2)^{1/2}\, dv = z^2 = \frac{\kappa\varepsilon}{\phi(0)} + O(\varepsilon^2),$$

and

$$\frac{2}{\pi}\int_{z_1}^{z_2}\{(z_2-x)(x-z_1)\}^{1/2}x\,dx=\frac{2}{\pi}\int_{-z}^{z}(z^2-v^2)^{1/2}\{\tfrac{1}{2}(z_1+z_2)+v\}\,dv$$
$$=\tfrac{1}{2}z^2(z_1+z_2)=O(\varepsilon^2).$$

Substitution of these results in (4.13) yields the required result (4.11).

LEMMA 4.3. *As* $\varepsilon\to 0$

$$\alpha^2=\kappa\varepsilon+\theta(\varepsilon,\kappa)\varepsilon^2+O(\varepsilon^{5/2}),\tag{4.16}$$

uniformly with respect to κ, *where* $\theta(\varepsilon,\kappa)$ *is continuous in the rectangle*

$$0\leqq\varepsilon\leqq\delta,\qquad k_1\leqq\kappa\leqq k_2.$$

To prove this result, we repeat the steps of the proof of Lemma 4.2, retaining more terms in the expansions. Thus in place of (4.12), we use

$$p(\varepsilon,\kappa,x)=\phi^2(0)+xp'(\varepsilon,\kappa,0)+\tfrac{1}{2}x^2p''(\varepsilon,\kappa,0)+O(\varepsilon^{3/2});$$

compare again Lemma 3.1. Also, the form of the next term in each of the expansions (3.04) is obtained by replacing the error terms $O(\varepsilon^{3/2})$ by $\mp Z(\kappa)\varepsilon^{3/2}+O(\varepsilon^2)$, where $Z(\kappa)$ is a continuous function of κ. Further details are straightforward, and need not be recorded.

LEMMA 4.4. *As* $\varepsilon\to 0$

$$\tfrac{1}{2}\zeta^2=|P(x)|+O(\varepsilon),\tag{4.17}$$

where $P(x)$ *is defined by* (2.08) *and the O-term is uniform with respect to* $\kappa\in[k_1,k_2]$ *and* $x\in[a,b]$.

Suppose first that $x\in[z_2,b]$. Then $\zeta\in[\alpha,\bar{b}]$, and the relation between x and ζ is expressed by (4.04). The right-hand side of this equation is evaluable in closed form; thus

$$\int_{\alpha}^{\zeta}(\tau^2-\alpha^2)^{1/2}\,d\tau=-\frac{1}{2}\alpha^2\ln\left(\frac{\zeta}{\alpha}\right)-\frac{1}{2}\alpha^2\ln\left\{1+\left(1-\frac{\alpha^2}{\zeta^2}\right)^{1/2}\right\}+\frac{1}{2}\zeta^2\left(1-\frac{\alpha^2}{\zeta^2}\right)^{1/2}.$$

Since $\zeta\geqq\alpha$, it follows that

$$\int_{\alpha}^{\zeta}(\tau^2-\alpha^2)^{1/2}\,d\tau=-\frac{1}{2}\alpha^2\ln\left(\frac{\zeta}{\alpha}\right)+\frac{1}{2}\zeta^2+O(\alpha^2)\tag{4.18}$$

as $\alpha\to 0$. Here and in the rest of the proof the O-symbol is understood to be uniform with respect to all variables.

To approximate the left-hand side of (4.04), we substitute by means of (3.05) and use Lemma 3.2. This yields

$$\int_{z_2}^{x}\{f(\varepsilon,\kappa,t)\}^{1/2}\,dt=\int_{z_2}^{x}\{(t-z_1)(t-z_2)\}^{1/2}\phi(t)\,dt+O(\varepsilon).$$

On making the substitutions (4.14), with x replaced by t, we obtain

$$\int_{z_2}^{x}\{f(\varepsilon,\kappa,t)\}^{1/2}\,dt=\int_{z}^{x-(1/2)(z_1+z_2)}(v^2-z^2)^{1/2}\phi\{\tfrac{1}{2}(z_1+z_2)+v\}\,dv+O(\varepsilon).$$

Since $z_1+z_2=O(\varepsilon)$, this gives

$$\int_{z_2}^{x}\{f(\varepsilon,\kappa,t)\}^{1/2}\,dt=\int_{z}^{x}(v^2-z^2)^{1/2}\phi(v)\,dv+O(\varepsilon)$$
$$=J_1+J_2-J_3+O(\varepsilon),$$

where

$$J_1=\int_z^x (v^2-z^2)^{1/2}\phi(0)\,dv,\qquad J_2=\int_z^x v\{\phi(v)-\phi(0)\}\,dv,$$
$$J_3=\int_z^x \{v-(v^2-z^2)^{1/2}\}\{\phi(v)-\phi(0)\}\,dv.$$

By comparing with (4.18) and recalling that $z=O(\varepsilon^{1/2})$, we see that

$$J_1=\phi(0)\left\{-\frac{1}{2}z^2\ln\left(\frac{x}{z}\right)+\frac{1}{2}x^2\right\}+O(\varepsilon).$$

Secondly,

$$J_2=\int_0^x v\{\phi(v)-\phi(0)\}\,dv+O(\varepsilon)=\int_0^x v\phi(v)\,dv-\frac{1}{2}\phi(0)x^2+O(\varepsilon).$$

Thirdly,

$$|J_3|\leqq z^2\int_z^x\frac{|\phi(v)-\phi(0)|}{v}\,dv\leqq z^2\int_0^x\frac{|\phi(v)-\phi(0)|}{v}\,dv=O(\varepsilon).$$

Combining these results we obtain

$$\int_{z_2}^{x}\{f(\varepsilon,\kappa,t)\}^{1/2}\,dt=\int_0^x v\phi(v)\,dv-\frac{1}{2}\phi(0)z^2\ln\left(\frac{x}{z}\right)+O(\varepsilon).$$

In consequence of (4.04), the right-hand side of this relation may be equated to the right-hand side of (4.18). On using Lemma 4.2 and (4.15), we then find that

$$\tfrac{1}{2}\zeta^2=\int_0^x v\phi(v)\,dv+\tfrac{1}{2}\kappa\varepsilon\ln\left(\frac{\zeta}{x}\right)+O(\varepsilon). \tag{4.19}$$

We now refer to the equation

$$\frac{dx}{d\zeta}=\left\{\frac{\zeta^2-\alpha^2}{(x-z_1)(x-z_2)}\right\}^{1/2}\frac{1}{\{p(\varepsilon,\kappa,x)\}^{1/2}}, \tag{4.20}$$

obtained from (3.05) and (4.01). From the proof of Lemma I of [12] used in establishing Lemma 4.1, it is immediately seen that the right-hand side of (4.20) is continuous and positive. Therefore $dx/d\zeta$ is bounded, and also bounded away from zero. By integration, it follows that $(x-z_2)/(\zeta-\alpha)$ and $(\zeta-\alpha)/(x-z_2)$ are both bounded. Using (3.04) and Lemma 4.2 we then see that x/ζ and ζ/x are both bounded, and hence that $|\ln(\zeta/x)|$ is bounded. Therefore from (4.19)

$$\tfrac{1}{2}\zeta^2=\int_0^x v\phi(v)\,dv+O(\varepsilon)=|P(x)|+O(\varepsilon);$$

compare (2.08) and (3.01).

We have therefore established the required result (4.17) when $x \in [z_2, b]$. An exactly similar proof holds when $x \in [a, z_1]$. Lastly, if $x \in [z_1, z_2]$, then $|\zeta| \leqq \alpha$ and hence from Lemma 4.2 $\frac{1}{2}\zeta^2$ is $O(\varepsilon)$. Also $P(x)$ is $O(x^2)$ and therefore $O(\varepsilon)$. Thus (4.17) again applies. This completes the proof of Lemma 4.4.

5. Uniform approximations to the solution of the boundary-value problem. Instead of applying the approximation theorems of [12] directly to the differential equation (4.07), we first modify this equation in the following manner. Write

$$\bar{\psi}(\varepsilon, \kappa, \zeta) = \psi(\alpha, \kappa, \zeta) + \varepsilon^{-2}(\kappa\varepsilon - \alpha^2). \tag{5.01}$$

Then (4.07) becomes

$$d^2W/d\zeta^2 = \{\varepsilon^{-2}(\zeta^2 - \kappa\varepsilon) + \bar{\psi}(\varepsilon, \kappa, \zeta)\}W, \tag{5.02}$$

and from Lemmas 4.1 and 4.3 it follows that $\bar{\psi}(\varepsilon, \kappa, \zeta)$ is a continuous function of its arguments.

Applying Theorem I of [12] to (5.02) and using the same notation as in this reference, we deduce that there are solutions $W_1(\zeta) \equiv W_1(\varepsilon, \kappa, \zeta)$ and $W_2(\zeta) \equiv W_2(\varepsilon, \kappa, \zeta)$ having the following forms in the interval $0 \leqq \zeta \leqq \bar{b}$:

$$W_1(\zeta) = U(\zeta\sqrt{2/\varepsilon}) + \eta_1(\zeta), \qquad W_2(\zeta) = \bar{U}(\zeta\sqrt{2/\varepsilon}) + \eta_2(\zeta), \tag{5.03}$$

where the error terms $\eta_1(\zeta) \equiv \eta_1(\varepsilon, \kappa, \zeta)$ and $\eta_2(\zeta) \equiv \eta_2(\varepsilon, \kappa, \zeta)$ are subject to the bounds

(5.04)

$$\frac{|\eta_1(\zeta)|}{\mathbf{M}(\zeta\sqrt{2/\varepsilon})}, \frac{|\eta_1'(\zeta)|}{(2/\varepsilon)^{1/2}\mathbf{N}(\zeta\sqrt{2/\varepsilon})} \leqq \mathbf{E}^{-1}(\zeta\sqrt{2/\varepsilon})[\exp\{\tfrac{1}{2}(\pi\varepsilon)^{1/2}l_1(-\tfrac{1}{2}\kappa)\mathcal{V}_{\zeta,\bar{b}}(F)\} - 1],$$

(5.05)

$$\frac{|\eta_2(\zeta)|}{\mathbf{M}(\zeta\sqrt{2/\varepsilon})}, \frac{|\eta_2'(\zeta)|}{(2/\varepsilon)^{1/2}\mathbf{N}(\zeta\sqrt{2/\varepsilon})} \leqq \mathbf{E}(\zeta\sqrt{2/\varepsilon})[\exp\{\tfrac{1}{2}(\pi\varepsilon)^{1/2}l_1(-\tfrac{1}{2}\kappa)\mathcal{V}_{0,\zeta}(F)\} - 1].$$

Here, and subsequently, $\eta_1'(\zeta)$ and $\eta_2'(\zeta)$ denote $\partial\eta_1/\partial\zeta$ and $\partial\eta_2/\partial\zeta$ respectively, and we have suppressed the argument parameter $-\frac{1}{2}\kappa$ of the parabolic cylinder functions U and $\bar{U}$ and also of the auxiliary functions $\mathbf{E}$, $\mathbf{M}$, and $\mathbf{N}$; thus

$$U(\zeta\sqrt{2/\varepsilon}) \equiv U(-\tfrac{1}{2}\kappa, \zeta\sqrt{2/\varepsilon}),$$

and so on. The error-control function $F \equiv F(\varepsilon, \kappa, \zeta)$ is defined by

$$F(\varepsilon, \kappa, \zeta) = \int \frac{\bar{\psi}(\varepsilon, \kappa, \zeta)}{\Omega(\zeta\sqrt{2/\varepsilon})}\, d\zeta, \tag{5.06}$$

with the balancing function $\Omega(t)$ subject to the conditions prescribed in [12, § 6.1].

On taking $\Omega(t)$ to be $1 + |t|$ and observing from Lemma 4.4 that $\bar{b}$ is bounded, we see that

$$\begin{aligned}\mathcal{V}_{0,\bar{b}}(F) &= \int_0^{\bar{b}} \frac{|\bar{\psi}(\varepsilon, \kappa, t)|}{1 + t\sqrt{2/\varepsilon}}\, dt \\ &= O(1)\int_0^{\bar{b}} \frac{dt}{1 + t\sqrt{2/\varepsilon}} = O\{\varepsilon^{1/2}\ln(1/\varepsilon)\}\end{aligned}$$

as $\varepsilon \to 0$, uniformly with respect to $\kappa \in [k_1, k_2]$. Next, the function $l_1(-\frac{1}{2}\kappa)$ is continuous in κ, and therefore bounded. Substituting in (5.04) and (5.05) by means of these results we conclude that

$$\begin{aligned}
&\eta_1(\zeta) = \mathbf{E}^{-1}(\zeta\sqrt{2/\varepsilon})\mathbf{M}(\zeta\sqrt{2/\varepsilon})O(\hat{\varepsilon}),\\
&\eta_1'(\zeta) = \varepsilon^{-1/2}\mathbf{E}^{-1}(\zeta\sqrt{2/\varepsilon})\mathbf{N}(\zeta\sqrt{2/\varepsilon})O(\hat{\varepsilon}),\\
&\eta_2(\zeta) = \mathbf{E}(\zeta\sqrt{2/\varepsilon})\mathbf{M}(\zeta\sqrt{2/\varepsilon})O(\hat{\varepsilon}),\\
&\eta_2'(\zeta) = \varepsilon^{-1/2}\mathbf{E}(\zeta\sqrt{2/\varepsilon})\mathbf{N}(\zeta\sqrt{2/\varepsilon})O(\hat{\varepsilon}),
\end{aligned} \tag{5.07}$$

as $\varepsilon \to 0$ uniformly with respect to $\kappa \in [k_1, k_2]$ and $\zeta \in [0, \bar{b}]$, where

$$\hat{\varepsilon} = \varepsilon \ln(1/\varepsilon). \tag{5.08}$$

We also note here the properties

$$\eta_1(\bar{b}) = \eta_1'(\bar{b}) = 0, \qquad \eta_2(0) = \eta_2'(0) = 0, \tag{5.09}$$

obtained from (5.04) and (5.05) by setting $\zeta = \bar{b}$ and 0, respectively.

In a similar manner, by applying Theorem I of [12] to (5.02) with ζ replaced by $-\zeta$, we deduce that in the interval $\bar{a} \leqq \zeta \leqq 0$ there are solutions $W_3(\zeta) \equiv W_3(\varepsilon, \kappa, \zeta)$ and $W_4(\zeta) \equiv W_4(\varepsilon, \kappa, \zeta)$, given by

$$W_3(\zeta) = U(-\zeta\sqrt{2/\varepsilon}) + \eta_3(\zeta), \qquad W_4(\zeta) = \bar{U}(-\zeta\sqrt{2/\varepsilon}) + \eta_4(\zeta), \tag{5.10}$$

where

$$\begin{aligned}
&\eta_3(\zeta) = \mathbf{E}^{-1}(-\zeta\sqrt{2/\varepsilon})\mathbf{M}(-\zeta\sqrt{2/\varepsilon})O(\hat{\varepsilon}),\\
&\eta_3'(\zeta) = \varepsilon^{-1/2}\mathbf{E}^{-1}(-\zeta\sqrt{2/\varepsilon})\mathbf{N}(-\zeta\sqrt{2/\varepsilon})O(\hat{\varepsilon}),\\
&\eta_4(\zeta) = \mathbf{E}(-\zeta\sqrt{2/\varepsilon})\mathbf{M}(-\zeta\sqrt{2/\varepsilon})O(\hat{\varepsilon}),\\
&\eta_4'(\zeta) = \varepsilon^{-1/2}\mathbf{E}(-\zeta\sqrt{2/\varepsilon})\mathbf{N}(-\zeta\sqrt{2/\varepsilon})O(\hat{\varepsilon}),
\end{aligned} \tag{5.11}$$

as $\varepsilon \to 0$ uniformly with respect to $\kappa \in [k_1, k_2]$ and $\zeta \in [\bar{a}, 0]$. Furthermore,

$$\eta_3(\bar{a}) = \eta_3'(\bar{a}) = 0, \qquad \eta_4(0) = \eta_4'(0) = 0. \tag{5.12}$$

To determine the boundary conditions for (5.02) we require the values of $\dot{x}^{-1/2}$ at the endpoints. Using (3.02), (4.01), Lemma 4.2, and Lemma 4.4, and recalling that $\phi(x) = |A_0(x)|$, we find that

$$\begin{aligned}
&\dot{x}^2 = 2\{aA_0(a)\}^{-2}|P(a)| + O(\varepsilon) \quad \text{at } x = a,\\
&\dot{x}^2 = 2\{bA_0(b)\}^{-2}|P(b)| + O(\varepsilon) \quad \text{at } x = b.
\end{aligned}$$

From these results and equations (2.10), (2.11), and (4.06) it follows that the wanted solution $W(\zeta) \equiv W(\varepsilon, \kappa, \zeta)$ of (5.02) satisfies the conditions

$$W(\bar{a}) = \bar{l}\, e^{P(a)/\varepsilon}(1+\lambda), \qquad W(\bar{b}) = \bar{m}\, e^{P(b)/\varepsilon}(1+\mu), \tag{5.13}$$

where $\bar{l}$ and $\bar{m}$ are constants given by

$$\bar{l} = |aA_0(a)|^{1/2}|2P(a)|^{-1/4}\, e^{Q(a)}l, \qquad \bar{m} = |bA_0(b)|^{1/2}|2P(b)|^{-1/4}\, e^{Q(b)}m, \tag{5.14}$$

and λ and μ are functions of ε and κ with the properties

$$\lambda = O(\varepsilon), \qquad \mu = O(\varepsilon), \tag{5.15}$$

as $\varepsilon \to 0$, uniformly with respect to κ. We now write

$$\begin{aligned} W(\zeta) &= \Lambda_1 W_1(\zeta) + \Lambda_2 W_2(\zeta), \qquad 0 \leqq \zeta \leqq \bar{b}, \\ W(\zeta) &= \Lambda_3 W_3(\zeta) + \Lambda_4 W_4(\zeta), \qquad \bar{a} \leqq \zeta \leqq 0, \end{aligned} \tag{5.16}$$

where the coefficients Λ_1, Λ_2, Λ_3, and Λ_4 are to be determined.

When $\zeta = 0$, the properties of the parabolic cylinder functions and the auxiliary functions given in [12, § 5] show that $\mathbf{E}(0) = 1$ and

$$\begin{aligned} U(0) &= \mathbf{M}(0) \sin\beta, \qquad U'(0) = -\mathbf{N}(0)\cos\beta, \\ \bar{U}(0) &= \mathbf{M}(0)\cos\beta, \qquad \bar{U}'(0) = \mathbf{N}(0)\sin\beta, \end{aligned} \tag{5.17}$$

where

$$\beta = \tfrac{1}{4}\pi(1+\kappa). \tag{5.18}$$

From (5.03), (5.09), (5.10), (5.12), and (5.17), we obtain

$$\begin{aligned} W_1(0) &= \mathbf{M}(0)\,(\sin\beta + \rho_1), & W_1'(0) &= -(2/\varepsilon)^{1/2}\mathbf{N}(0)(\cos\beta + \sigma_1), \\ W_2(0) &= \mathbf{M}(0)\cos\beta, & W_2'(0) &= (2/\varepsilon)^{1/2}\mathbf{N}(0)\sin\beta, \\ W_3(0) &= \mathbf{M}(0)(\sin\beta + \rho_3), & W_3'(0) &= (2/\varepsilon)^{1/2}\mathbf{N}(0)(\cos\beta + \sigma_3), \\ W_4(0) &= \mathbf{M}(0)\cos\beta, & W_4'(0) &= -(2/\varepsilon)^{1/2}\mathbf{N}(0)\sin\beta, \end{aligned}$$

where

$$\begin{aligned} \rho_1 &= \frac{\eta_1(0)}{\mathbf{M}(0)}, \qquad \sigma_1 = -\left(\frac{\varepsilon}{2}\right)^{1/2}\frac{\eta_1'(0)}{\mathbf{N}(0)}, \\ \rho_3 &= \frac{\eta_3(0)}{\mathbf{M}(0)}, \qquad \sigma_3 = \left(\frac{\varepsilon}{2}\right)^{1/2}\frac{\eta_3'(0)}{\mathbf{N}(0)}. \end{aligned} \tag{5.19}$$

Thus ρ_1, σ_1, ρ_3, and σ_3 are functions of ε and κ, and from (5.07) and (5.11) it follows that

$$\rho_1, \sigma_1, \rho_3, \sigma_3 = O(\hat{\varepsilon}) \tag{5.20}$$

as $\varepsilon \to 0$, uniformly with respect to κ.

On referring to (5.16) and matching the solutions and their ζ-derivatives at $\zeta = 0$, we obtain

$$\begin{aligned} \Lambda_1(\sin\beta + \rho_1) + \Lambda_2\cos\beta &= \Lambda_3(\sin\beta + \rho_3) + \Lambda_4\cos\beta, \\ -\Lambda_1(\cos\beta + \sigma_1) + \Lambda_2\sin\beta &= \Lambda_3(\cos\beta + \sigma_3) - \Lambda_4\sin\beta. \end{aligned}$$

The boundary conditions (5.13) yield

$$\Lambda_3 U(-\bar{a}\sqrt{2/\varepsilon}) + \Lambda_4\bar{U}(-\bar{a}\sqrt{2/\varepsilon})(1+\nu_4) = \bar{l}\, e^{P(a)/\varepsilon}(1+\lambda),$$

$$\Lambda_1 U(\bar{b}\sqrt{2/\varepsilon}) + \Lambda_2\bar{U}(\bar{b}\sqrt{2/\varepsilon})(1+\nu_2) = \bar{m}\, e^{P(b)/\varepsilon}(1+\mu),$$

where ν_2 and ν_4 denote the quantities

$$\nu_2 = \eta_2(\bar{b})/\bar{U}(\bar{b}\sqrt{2/\varepsilon}), \qquad \nu_4 = \eta_4(\bar{a})/\bar{U}(-\bar{a}\sqrt{2/\varepsilon}). \tag{5.21}$$

From (5.07), (5.11), and the properties of the auxiliary functions **E** and **M** given in [12, § 5.8], we see that

(5.22) $$\nu_2,\ \nu_4 = O(\hat{\varepsilon})$$

as $\varepsilon \to 0$, uniformly with respect to κ.

Solving the four simultaneous equations just given for Λ_1, Λ_2, Λ_3, and Λ_4, we obtain, after a straightforward calculation,

(5.23) $$D\Lambda_1 = \bar{m}\, e^{P(b)/\varepsilon}(1+\mu)\sin 2\beta U(-\bar{a}\sqrt{2/\varepsilon}) + \bar{m}\, e^{P(b)/\varepsilon}(1+\mu)(\cos 2\beta + \sigma_3\cos\beta - \rho_3\sin\beta)(1+\nu_4)\bar{U}(-\bar{a}\sqrt{2/\varepsilon}) - \bar{l}\, e^{P(a)/\varepsilon}(1+\lambda)(1+\sigma_3\cos\beta+\rho_3\sin\beta)(1+\nu_2)\bar{U}(\bar{b}\sqrt{2/\varepsilon}),$$

(5.24) $$D\Lambda_2 = \bar{l}\, e^{P(a)/\varepsilon}(1+\lambda)(1+\sigma_3\cos\beta+\rho_3\sin\beta)U(\bar{b}\sqrt{2/\varepsilon}) + \bar{m}\, e^{P(b)/\varepsilon}(1+\mu)(\cos 2\beta+\sigma_1\cos\beta-\rho_1\sin\beta)U(-\bar{a}\sqrt{2/\varepsilon}) - \bar{m}\, e^{P(b)/\varepsilon}(1+\mu) \cdot \{\sin 2\beta + (\rho_1+\rho_3)\cos\beta + (\sigma_1+\sigma_3)\sin\beta + \rho_1\sigma_3 + \rho_3\sigma_1\} \cdot (1+\nu_4)\bar{U}(-\bar{a}\sqrt{2/\varepsilon}),$$

(5.25) $$D\Lambda_3 = \bar{l}\, e^{P(a)/\varepsilon}(1+\lambda)\sin 2\beta U(\bar{b}\sqrt{2/\varepsilon}) + \bar{l}\, e^{P(a)/\varepsilon}(1+\lambda)(\cos 2\beta+\sigma_1\cos\beta-\rho_1\sin\beta)(1+\nu_2)\bar{U}(\bar{b}\sqrt{2/\varepsilon}) - \bar{m}\, e^{P(b)/\varepsilon}(1+\mu)(1+\sigma_1\cos\beta+\rho_1\sin\beta)(1+\nu_4)\bar{U}(-\bar{a}\sqrt{2/\varepsilon}),$$

(5.26) $$D\Lambda_4 = \bar{m}\, e^{P(b)/\varepsilon}(1+\mu)(1+\sigma_1\cos\beta+\rho_1\sin\beta)U(-\bar{a}\sqrt{2/\varepsilon}) + \bar{l}\, e^{P(a)/\varepsilon}(1+\lambda)(\cos 2\beta+\sigma_3\cos\beta-\rho_3\sin\beta)U(\bar{b}\sqrt{2/\varepsilon}) - \bar{l}\, e^{P(a)/\varepsilon}(1+\lambda) \cdot \{\sin 2\beta + (\rho_1+\rho_3)\cos\beta + (\sigma_1+\sigma_3)\sin\beta + \rho_1\sigma_3 + \rho_3\sigma_1\} \cdot (1+\nu_2)\bar{U}(\bar{b}\sqrt{2/\varepsilon}),$$

where

(5.27) $$D = \sin 2\beta U(-\bar{a}\sqrt{2/\varepsilon})U(\bar{b}\sqrt{2/\varepsilon}) + (\cos 2\beta+\sigma_1\cos\beta-\rho_1\sin\beta)(1+\nu_2)U(-\bar{a}\sqrt{2/\varepsilon})\bar{U}(\bar{b}\sqrt{2/\varepsilon}) + (\cos 2\beta+\sigma_3\cos\beta-\rho_3\sin\beta)(1+\nu_4)\bar{U}(-\bar{a}\sqrt{2/\varepsilon})U(\bar{b}\sqrt{2/\varepsilon}) - \{\sin 2\beta + (\rho_1+\rho_3)\cos\beta + (\sigma_1+\sigma_3)\sin\beta + \rho_1\sigma_3 + \rho_3\sigma_1\} \cdot (1+\nu_2)(1+\nu_4)\bar{U}(-\bar{a}\sqrt{2/\varepsilon})\bar{U}(\bar{b}\sqrt{2/\varepsilon}).$$

These results supply the desired formulas for the coefficients Λ_1, Λ_2, Λ_3, and Λ_4 in the expressions (5.16) for the wanted solution $W(\zeta)$. This solution is related to that of the original boundary-value problem of § 2 via

$$y(x)=\exp\left\{-\frac{1}{\varepsilon}\int_0^x tA(\varepsilon,t)\,dt\right\}\dot{x}^{1/2}W(\zeta); \tag{5.28}$$

compare (2.03) and (4.06). Hence from (2.09) and (5.16) we have

$$y(x)=\{1+O(\varepsilon)\}\dot{x}^{1/2}e^{-Q(x)}e^{-P(x)/\varepsilon}\{\Lambda_1W_1(\zeta)+\Lambda_2W_2(\zeta)\},\qquad 0\leqq x\leqq b,$$

and

$$y(x)=\{1+O(\varepsilon)\}\dot{x}^{1/2}e^{-Q(x)}e^{-P(x)/\varepsilon}\{\Lambda_3W_3(\zeta)+\Lambda_4W_4(\zeta)\},\qquad a\leqq x\leqq 0,$$

where the O-terms are uniform with respect to κ and x. In the proof of Lemma 4.4 we noted that $\dot{x}$ and its reciprocal are continuous and bounded for all values of the variables. In consequence, *as $\varepsilon\to 0$ the asymptotic rate of growth (or decay) of the solution $y(x)$ is governed entirely by that of the functions*

$$\Lambda_1e^{-P(x)/\varepsilon}W_1(\zeta)+\Lambda_2e^{-P(x)/\varepsilon}W_2(\zeta) \tag{5.29}$$

when $0\leqq x\leqq b$, or

$$\Lambda_3e^{-P(x)/\varepsilon}W_3(\zeta)+\Lambda_4e^{-P(x)/\varepsilon}W_4(\zeta) \tag{5.30}$$

when $a\leqq x\leqq 0$.

Investigation of this growth forms the subject of the next three sections. Because we shall have frequent need of asymptotic estimates for the parabolic cylinder functions, we collect the relevant results here for reference.

By combining Lemma 4.4 with the asymptotic forms of the parabolic cylinder functions of large positive argument given in [10, Chap. 19] or [12, § 5.2], and remembering that the argument parameter of each of these functions is $-\frac{1}{2}\kappa$, we derive

$$\begin{aligned}U(|\zeta|\sqrt{2/\varepsilon})&=\varepsilon^{(1-\kappa)/4}e^{-|P(x)|/\varepsilon}e^{O(1)},\\ \bar{U}(|\zeta|\sqrt{2/\varepsilon})&=\varepsilon^{(1+\kappa)/4}e^{|P(x)|/\varepsilon}e^{O(1)},\end{aligned} \tag{5.31}$$

as $\varepsilon\to 0$, uniformly with respect to κ and x, *provided that $\kappa\in[k_1,k_2]$ and $|x|$ is bounded away from zero.* In the same circumstances,

$$\begin{aligned}\mathbf{E}^{-1}(|\zeta|\sqrt{2/\varepsilon})\mathbf{M}(|\zeta|\sqrt{2/\varepsilon})&=\varepsilon^{(1-\kappa)/4}e^{-|P(x)|/\varepsilon}e^{O(1)},\\ \mathbf{E}(|\zeta|\sqrt{2/\varepsilon})\mathbf{M}(|\zeta|\sqrt{2/\varepsilon})&=\varepsilon^{(1+\kappa)/4}e^{|P(x)|/\varepsilon}e^{O(1)}.\end{aligned} \tag{5.32}$$

Particular cases of (5.31) are given by

$$\begin{aligned}U(-\bar{a}\sqrt{2/\varepsilon})&=\varepsilon^{(1-\kappa)/4}e^{-|P(a)|/\varepsilon}e^{O(1)},\\ \bar{U}(-\bar{a}\sqrt{2/\varepsilon})&=\varepsilon^{(1+\kappa)/4}e^{|P(a)|/\varepsilon}e^{O(1)},\end{aligned} \tag{5.33}$$

$$\begin{aligned}U(\bar{b}\sqrt{2/\varepsilon})&=\varepsilon^{(1-\kappa)/4}e^{-|P(b)|/\varepsilon}e^{O(1)},\\ \bar{U}(\bar{b}\sqrt{2/\varepsilon})&=\varepsilon^{(1+\kappa)/4}e^{|P(b)|/\varepsilon}e^{O(1)},\end{aligned} \tag{5.34}$$

valid when $\varepsilon\to 0$, uniformly with respect to κ.

For the full interval $x \in [a, b]$, we have the weaker estimates

$$U(|\zeta|\sqrt{2/\varepsilon}) = \varepsilon^{(1-\kappa_1)/4} e^{-|P(x)|/\varepsilon} O(1), \tag{5.35}$$
$$\bar{U}(|\zeta|\sqrt{2/\varepsilon}) = e^{|P(x)|/\varepsilon} O(1),$$

$$\mathbf{E}^{-1}(|\zeta|\sqrt{2/\varepsilon})\mathbf{M}(|\zeta|\sqrt{2/\varepsilon}) = \varepsilon^{(1-\kappa_1)/4} e^{-|P(x)|/\varepsilon} O(1), \tag{5.36}$$
$$\mathbf{E}(|\zeta|\sqrt{2/\varepsilon})\mathbf{M}(|\zeta|\sqrt{2/\varepsilon}) = e^{|P(x)|/\varepsilon} O(1),$$

valid when $\varepsilon \to 0$, uniformly with respect to κ and x, where

$$\kappa_1 = \max(\kappa, 1). \tag{5.37}$$

6. Conditions for resonance: Case $A_0(x)$ negative. In this section we use the expressions (5.29) and (5.30) to investigate the asymptotic behavior of the solution of the boundary-value problem of § 2 in the case when the coefficient $A_0(x)$ in (2.01) is negative throughout the given x-interval $[a, b]$. From (2.08) we observe that this condition implies that the function $P(x)$ is negative for all values of x, other than $x = 0$.

Suppose first that the positive parameter κ is independent of ε and not an odd integer. From (5.18) and (2.02) it is seen that this implies both β and $B(0, 0)$ are fixed, $\sin 2\beta \neq 0$, and

$$B(0, 0) \neq 0, 2, 4, \cdots. \tag{6.01}$$

Referring to (5.20), (5.22), (5.33), and (5.34), we observe that as $\varepsilon \to 0$ the dominant contribution to D in (5.27) derives from the last line; consequently

$$\frac{1}{D} = \frac{O(1)}{\bar{U}(-\bar{a}\sqrt{2/\varepsilon})\bar{U}(\bar{b}\sqrt{2/\varepsilon})}. \tag{6.02}$$

Next, on substituting in (5.23) by means of (5.15), (5.20), and (5.22) and then referring to (5.33) we see that

$$D\Lambda_1 = O(1)\, e^{P(b)/\varepsilon}\bar{U}(-\bar{a}\sqrt{2/\varepsilon}) + O(1)\, e^{P(a)/\varepsilon}\bar{U}(\bar{b}\sqrt{2/\varepsilon}). \tag{6.03}$$

Also, from (5.03) and (5.07) we obtain

$$W_1(\zeta) = O(1)\mathbf{E}^{-1}(\zeta\sqrt{2/\varepsilon})\mathbf{M}(\zeta\sqrt{2/\varepsilon}).$$

Combining the last three equations, we see that when $0 \leqq x \leqq b$

$$\Lambda_1 e^{-P(x)/\varepsilon} W_1(\zeta) = O(1)\frac{e^{\{P(b)-P(x)\}/\varepsilon}}{\bar{U}(\bar{b}\sqrt{2/\varepsilon})}\mathbf{E}^{-1}(\zeta\sqrt{2/\varepsilon})\mathbf{M}(\zeta\sqrt{2/\varepsilon}) \tag{6.04}$$
$$+ O(1)\frac{e^{\{P(a)-P(x)\}/\varepsilon}}{\bar{U}(-\bar{a}\sqrt{2/\varepsilon})}\mathbf{E}^{-1}(\zeta\sqrt{2/\varepsilon})\mathbf{M}(\zeta\sqrt{2/\varepsilon}).$$

Using (5.33), (5.34), and (5.36), and bearing in mind that $P(x) = -|P(x)|$ throughout the present section, we find that

$$\Lambda_1 e^{-P(x)/\varepsilon} W_1(\zeta) = O(1)\varepsilon^{-(\kappa+\kappa_1)/4}\{e^{-2|P(b)|/\varepsilon} + e^{-2|P(a)|/\varepsilon}\}. \tag{6.05}$$

This vanishes as $\varepsilon \to 0$, uniformly with respect to x.

Next, by similar analysis we obtain from (5.03), (5.07), (5.24), and (6.02)

$$\Lambda_2\, e^{-P(x)/\varepsilon} W_2(\zeta) = O(1) \frac{e^{\{P(b)-P(x)\}/\varepsilon}}{\bar{U}(\bar{b}\sqrt{2/\varepsilon})} \mathbf{E}(\zeta\sqrt{2/\varepsilon})\mathbf{M}(\zeta\sqrt{2/\varepsilon}), \tag{6.06}$$

and hence

$$\Lambda_2\, e^{-P(x)/\varepsilon} W_2(\zeta) = O(1)\varepsilon^{-(1+\kappa)/4}\, e^{2\{|P(x)|-|P(b)|\}/\varepsilon}. \tag{6.07}$$

Let b_1 be any constant such that $0<b_1<b$. Because $|P(x)|$ is increasing when x is positive, the right-hand side of (6.07) vanishes as $\varepsilon \to 0$, uniformly with respect to $x \in [0, b_1]$.

In a similar manner it is verifiable that $\Lambda_3\, e^{-P(x)/\varepsilon} W_3(\zeta)$ vanishes uniformly in $[a, 0]$, and $\Lambda_4\, e^{-P(x)/\varepsilon} W_4(\zeta)$ vanishes uniformly in $[a_1, 0]$, where a_1 is any constant such that $a<a_1<0$.

On combining the foregoing results we conclude that when κ is independent of ε and not an odd integer, the solution $y(x)$ of the boundary-value problem of § 2 vanishes as $\varepsilon \to 0$, uniformly in any closed interval properly interior to the given interval $[a, b]$. That is, there is no resonance in the sense defined in § 1. This conclusion is already well known from the theory of Ackerberg and O'Malley, given in [1], and also [14, Chap. 8].

In order to induce resonance it is clear that as a first step we must arrange for the contribution from the term in $\bar{U}(-\bar{a}\sqrt{2/\varepsilon})\bar{U}(\bar{b}\sqrt{2/\varepsilon})$ in the expression (5.27) for D not to dominate the other terms. This is achieved by imposing the condition[2]

$$\sin 2\beta + (\rho_1+\rho_3)\cos\beta + (\sigma_1+\sigma_3)\sin\beta + \rho_1\sigma_3 + \rho_3\sigma_1 = O(e^{-2\varpi/\varepsilon}), \tag{6.08}$$

where ϖ is any constant such that

$$\varpi > \max\{|P(a)|, |P(b)|\}. \tag{6.09}$$

The condition (6.08) is satisfied by relaxing the requirement that β and κ be independent of ε. Thus (6.08) is to be regarded as an equation for $\beta=\beta(\varepsilon)$, say, and the corresponding value of $\kappa=\kappa(\varepsilon)$ is given by (5.18). From (5.20) we see that for small values of ε the relevant roots are

$$\beta = \tfrac{1}{2}s\pi + O(\hat{\varepsilon}), \tag{6.10}$$

where $s = 1, 2, 3, \cdots$ and $\hat{\varepsilon}$ is given by (5.08). Correspondingly,

$$\kappa = 2s - 1 + O(\hat{\varepsilon}). \tag{6.11}$$

We note that for each value of s, neither the value of β nor that of κ is unique because of the freedom of choice allowed by inclusion of the O-term on the right-hand side of (6.08). All roots satisfy (6.10), however.

We now suppose that $\beta=\beta(\varepsilon)$ is prescribed in the manner just indicated, and distinguish three cases in subsequent analysis, given by $|P(a)| \gtreqless |P(b)|$. In examining these cases in turn, we continue to use the symbols a_1 and b_1 to denote any positive constants such that $a<a_1<0$ and $0<b_1<b$; in addition we use a_2 and b_2 to denote any constants such that $a_1<a_2<0$ and $0<b_2<b_1$.

Case I. $|P(a)|>|P(b)|$. In this case Lemma 4.4 shows that $|\bar{a}|>\bar{b}$ for all sufficiently small values of ε, and referring to (5.33), (5.34), (6.08), and (6.09) and substituting in (5.27) by means of (5.20), (5.22), and (6.10), we see that

$$D = (-)^s\{1+O(\hat{\varepsilon})\}\bar{U}(-\bar{a}\sqrt{2/\varepsilon})U(\bar{b}\sqrt{2/\varepsilon}). \tag{6.12}$$

---

[2] The condition (6.08) is adequate for the present purposes, but it could be weakened somewhat without affecting the general conclusions.

From (5.23) we derive in a similar manner

$$D\Lambda_1=(-)^s\{1+O(\hat\varepsilon)\}\bar m\, e^{P(b)/\varepsilon}\bar U(-\bar a\sqrt{2/\varepsilon})-\{1+O(\hat\varepsilon)\}\bar l\, e^{P(a)/\varepsilon}\bar U(\bar b\sqrt{2/\varepsilon}). \tag{6.13}$$

Therefore

$$\begin{aligned}\Lambda_1\, e^{-P(x)/\varepsilon}W_1(\zeta)&=\{1+O(\hat\varepsilon)\}\bar m\frac{e^{\{P(b)-P(x)\}/\varepsilon}}{U(\bar b\sqrt{2/\varepsilon})}\{U(\zeta\sqrt{2/\varepsilon})+\eta_1(\zeta)\}\\&\quad+(-)^{s-1}\{1+O(\hat\varepsilon)\}\bar l\frac{e^{\{P(a)-P(x)\}/\varepsilon}\bar U(\bar b\sqrt{2/\varepsilon})}{\bar U(-\bar a\sqrt{2/\varepsilon})U(\bar b\sqrt{2/\varepsilon})}\{U(\zeta\sqrt{2/\varepsilon})+\eta_1(\zeta)\}.\end{aligned} \tag{6.14}$$

From (5.07), and also (5.33) through (5.36), the contribution from the second line is estimated by

$$O(1)\varepsilon^{(\kappa-\kappa_1)/4}\, e^{\{2|P(b)|-2|P(a)|\}/\varepsilon},$$

as $\varepsilon\to 0$, and therefore vanishes uniformly in $[0, b]$. On the other hand if $\bar m\neq 0$, that is (from (5.14)) if $m\neq 0$, then with the aid of (5.31) and (5.32) we see that in $[b_2, b]$ the contribution from the previous line is $e^{O(1)}$, that is, bounded and bounded away from zero.

Next, from (5.24), (6.08), and (6.10) we derive

$$\begin{aligned}D\Lambda_2&=\{1+O(\hat\varepsilon)\}\bar l e^{P(a)/\varepsilon}U(\bar b\sqrt{2/\varepsilon})+(-)^s\{1+O(\hat\varepsilon)\}\bar m\, e^{P(b)/\varepsilon}U(-\bar a\sqrt{2/\varepsilon})\\&\quad+O(e^{-2\varpi/\varepsilon})\, e^{P(b)/\varepsilon}\bar U(-\bar a\sqrt{2/\varepsilon}).\end{aligned} \tag{6.15}$$

Accordingly, from (5.33), (5.34), (6.09), and (6.12) it follows that

$$\Lambda_2=O(1)\, e^{P(a)/\varepsilon}/\bar U(-\bar a\sqrt{2/\varepsilon}).$$

Hence

$$\Lambda_2\, e^{-P(x)/\varepsilon}W_2(\zeta)=O(1)\frac{e^{\{P(a)-P(x)\}/\varepsilon}}{\bar U(-\bar a\sqrt{2/\varepsilon})}\mathbf{E}(\zeta\sqrt{2/\varepsilon})\mathbf{M}(\zeta\sqrt{2/\varepsilon}); \tag{6.16}$$

consequently

$$\Lambda_2\, e^{-P(x)/\varepsilon}W_2(\zeta)=O(1)\varepsilon^{-(1+\kappa)/4}\, e^{2\{|P(x)|-|P(a)|\}/\varepsilon}. \tag{6.17}$$

Since $|P(x)|\leqq|P(b)|$ this vanishes as $\varepsilon\to 0$, uniformly in $[0, b]$.

The behavior of the wanted solution in the interval $[a, 0]$ may be investigated in a similar manner. First, we find that as $\varepsilon\to 0$ the term $\Lambda_3\, e^{-P(x)/\varepsilon}W_3(\zeta)$ is bounded and bounded away from zero throughout $[a, a_2]$, unless $m=0$ in which event this term tends to zero uniformly in $[a, 0]$. Secondly, we find that $\Lambda_4\, e^{-P(x)/\varepsilon}W_4(\zeta)$ tends to zero uniformly in $[a_1, 0]$.

Combining the results for the intervals $[0, b]$ and $[a, 0]$ we reach the following conclusion: *when* $|P(a)|>|P(b)|$ *there is resonance if* $m\neq 0$ *but not if* $m=0$.

Case II. $|P(b)|>|P(a)|$. This case may be investigated in a similar manner to Case I, or the results can be deduced by symmetry. The conclusion is that *when* $|P(b)|>|P(a)|$ *there is resonance if* $l\neq 0$, *but not if* $l=0$.

Case III. $P(a)=P(b)$. In order to discuss this case we shall suppose that we know the value of $\gamma(a)-\gamma(b)$, where $\gamma(x)$ denotes the second coefficient in the expansion

$$\tag{6.18} \tfrac{1}{2}\zeta^2=|P(x)|+\gamma(x)\varepsilon+o(\varepsilon), \qquad \varepsilon\to 0;$$

compare Lemma 4.4. The value of $\gamma(a)-\gamma(b)$ can be found by extending the analysis of §§ 3 and 4, but we shall not pursue these details in the present paper. In passing, however, we note one special case in which the value of $\gamma(a)-\gamma(b)$ is already available. This is the case in which the coefficients $A(\varepsilon, x)$ and $B(\varepsilon, x)$ in (1.01) are even functions of x, and $a=-b$. In these circumstances the functions $f(\varepsilon, \kappa, x)$ and $P(x)$ are even in x, and $\zeta\equiv\zeta(x)$ is odd; in consequence $\gamma(a)-\gamma(b)=0$.

Returning to the general case, we see from Lemma 4.4 that when $P(a)=P(b)$ we have

$$\bar{a}=-\{1+O(\varepsilon)\}\bar{b}.$$

From this result, (6.18), and the asymptotic approximations for the parabolic cylinder functions of large positive argument, it follows that

$$\tag{6.19} U(-\bar{a}\sqrt{2/\varepsilon})=\{1+o(1)\}\, e^{\gamma(b)-\gamma(a)}U(\bar{b}\sqrt{2/\varepsilon}),$$

$$\tag{6.20} \bar{U}(-\bar{a}\sqrt{2/\varepsilon})=\{1+o(1)\}\, e^{\gamma(a)-\gamma(b)}\bar{U}(\bar{b}\sqrt{2/\varepsilon}).$$

Substituting in (5.27) by means of these approximations and using (6.08) and (6.10), we find that

$$\tag{6.21} D=(-)^s\{1+o(1)\}2\cosh\{\gamma(a)-\gamma(b)\}U(\bar{b}\sqrt{2/\varepsilon})\bar{U}(\bar{b}\sqrt{2/\varepsilon}).$$

The relation (6.13) remains valid in the present circumstances, and by referring to (6.20) we see that

$$\tag{6.22} D\Lambda_1=[(-)^s\{1+o(1)\}\bar{m}\, e^{\gamma(a)-\gamma(b)}-\{1+O(\hat{\varepsilon})\}\bar{l}\,]\, e^{P(b)/\varepsilon}\bar{U}(\bar{b}\sqrt{2/\varepsilon}).$$

We now impose the restriction

$$\tag{6.23} \bar{l}\neq(-)^s\bar{m}\, e^{\gamma(a)-\gamma(b)},$$

that is,

$$\tag{6.24} l\, e^{Q(a)-\gamma(a)}|aA_0(a)|^{1/2}\neq(-)^s m\, e^{Q(b)-\gamma(b)}|bA_0(b)|^{1/2};$$

compare (5.14). From (5.03), (6.21), and (6.22) we then obtain

$$\tag{6.25} \begin{aligned}&\Lambda_1\, e^{-P(x)/\varepsilon}W_1(\zeta)\\ &=\{1+o(1)\}\frac{\bar{m}\, e^{\gamma(a)-\gamma(b)}-(-)^s\bar{l}}{2\cosh\{\gamma(a)-\gamma(b)\}}\frac{e^{\{P(b)-P(x)\}/\varepsilon}}{U(\bar{b}\sqrt{2/\varepsilon})}\{U(\zeta\sqrt{2/\varepsilon})+\eta_1(\zeta)\}.\end{aligned}$$

As $\varepsilon\to 0$ this quantity is uniformly $e^{O(1)}$ in $[b_2, b]$.

Next, equation (6.15) remains valid. With the aid of (6.21) we deduce that

$$\tag{6.26} \Lambda_2=O(1)\, e^{P(b)/\varepsilon}/\bar{U}(\bar{b}\sqrt{2/\varepsilon}).$$

Accordingly,

$$\Lambda_2\, e^{-P(x)/\varepsilon}W_2(\zeta)=O(1)\varepsilon^{-(1+\kappa)/4}\, e^{2\{|P(x)|-|P(b)|\}/\varepsilon};$$

compare (6.17). This vanishes as $\varepsilon\to 0$, uniformly in $[0, b_1]$.

The analysis for the interval $[a, 0]$ is similar, and we find that as $\varepsilon \to 0$ the term $\Lambda_3 e^{-P(x)/\varepsilon} W_3(\zeta)$ is $e^{O(1)}$ uniformly in $[a, a_2]$, and the term $\Lambda_4 e^{-P(x)/\varepsilon} W_4(\zeta)$ vanishes uniformly in $[a_1, 0]$.

On combining the foregoing results we conclude that *when* $P(a) = P(b)$, *resonance takes place, provided that the condition* (6.23) (*or* (6.24)) *is satisfied.*

The problem is more difficult when

$$\bar{l} = (-)^s \bar{m}\, e^{\gamma(a)-\gamma(b)}. \tag{6.27}$$

The explicit terms on the right-hand side of (6.22) then cancel, but in general our analysis does not reveal whether or not the error terms cancel. There is, however, one special case in which we can arrive at a firm conclusion, as follows.

Let us suppose that the coefficients $A(\varepsilon, x)$ and $B(\varepsilon, x)$ in (1.01) are even in x, and $a = -b$. As noted above, the functions $P(x)$ and $\zeta \equiv \zeta(x)$ are then respectively even and odd, and $\gamma(a) = \gamma(b)$. It is also easily seen that $\bar{\psi}(\varepsilon, \kappa, \zeta)$ is an even function of ζ. From these properties it follows that the error terms are related by $\rho_1 = \rho_3$, $\sigma_1 = \sigma_3$, $\nu_2 = \nu_4$, and $\lambda = \mu$. Hence (6.08) reduces to

$$\sin 2\beta + 2\rho_3 \cos\beta + 2\sigma_3 \sin\beta + 2\rho_3\sigma_3 = O(e^{-2\varpi/\varepsilon}).$$

The left-hand side of this equation factors, and we deduce that either

$$\sin\beta + \rho_3 = O(e^{-2\varpi/\varepsilon}), \tag{6.28}$$

or

$$\cos\beta + \sigma_3 = O(e^{-2\varpi/\varepsilon}). \tag{6.29}$$

From (6.10) it is clear that the former alternative applies when s is even, and the latter when s is odd. In the present case the condition (6.27) reduces to $\bar{l} = (-)^s \bar{m}$. Whether s be even or odd, it is easily verified from (6.28) and (6.29) that

$$\bar{m}(\cos 2\beta + \sigma_3 \cos\beta - \rho_3 \sin\beta) - \bar{l}(1 + \sigma_3 \cos\beta + \rho_3 \sin\beta) = O(e^{-2\varpi/\varepsilon}).$$

Hence from (5.23) we obtain

$$D\Lambda_1 = \bar{m}\, e^{P(b)/\varepsilon}(1+\mu) \sin 2\beta\, U(\bar{b}\sqrt{2/\varepsilon}) + O(e^{-2\varpi/\varepsilon})\, e^{P(b)/\varepsilon}\, \bar{U}(\bar{b}\sqrt{2/\varepsilon}).$$

In consequence

$$\Lambda_1 e^{-P(x)/\varepsilon} W_1(\zeta) = O(\hat{\varepsilon}) \frac{e^{\{|P(x)|-|P(b)|\}/\varepsilon}}{\bar{U}(\bar{b}\sqrt{2/\varepsilon})} \mathbf{E}^{-1}(\zeta\sqrt{2/\varepsilon}) \mathbf{M}(\zeta\sqrt{2/\varepsilon}).$$

This vanishes as $\varepsilon \to 0$, uniformly in $[0, b]$. Similarly $\Lambda_3 e^{-P(x)/\varepsilon} W_3(\zeta)$ vanishes uniformly in $[a, 0]$. Accordingly, in this special case resonance does *not* take place.

7. Conditions for resonance: Case $A_0(x)$ positive. The analysis in this section parallels that of § 6, the essential change being that the function $P(x)$ is now nonnegative. We again use the symbols a_1, a_2, b_1, and b_2 to denote any constants such that $a < a_1 < a_2 < 0$ and $0 < b_2 < b_1 < b$.

Suppose first that the parameter κ is independent of ε and not an odd integer. From (5.18) and (2.02) it is seen that this implies both β and $B(0, 0)$ are fixed, $\sin 2\beta \neq 0$, and

$$B(0, 0) \neq -2, -4, -6, \cdots. \tag{7.01}$$

Equations (6.02), (6.04), and (6.06) again apply, and using (5.32), (5.33), (5.34), (5.36), and the fact that $P(x)$ is increasing when x is positive, we may verify that as $\varepsilon \to 0$, each of the quantities $\Lambda_1 e^{-P(x)/\varepsilon} W_1(\zeta)$ and $\Lambda_2 e^{-P(x)/\varepsilon} W_2(\zeta)$ is uniformly $O(\varepsilon^{-(\kappa+\kappa_1)/4})$ in $[0, b_2]$ and uniformly $O(1)$ in $[b_2, b]$ where κ_1 is given by (5.37). Similarly each of the quantities $\Lambda_3 e^{-P(x)/\varepsilon} W_3(\zeta)$ and $\Lambda_4 e^{-P(x)/\varepsilon} W_4(\zeta)$ is uniformly $O(\varepsilon^{-(\kappa+\kappa_1)/4})$ in $[a_2, 0]$ and uniformly $O(1)$ in $[a, a_2]$. These results accord with those of O'Malley [13, p. 487], and correspond to the nonresonant situation.

In the rest of this section, we again suppose that $\beta = \beta(\varepsilon)$ is a root of the equation (6.08), where ϖ is a constant subject to (6.09). Thus β and κ are estimated by (6.10) and (6.11), respectively.

Case I. $P(a) > P(b)$. Equation (6.14) again applies, and as $\varepsilon \to 0$ both terms on the right-hand side grow uniformly (and exponentially) in $[b_2, b_1]$. Equation (6.16) also holds, and from it we deduce that $\Lambda_2 e^{-P(x)/\varepsilon} W_2(\zeta)$ is uniformly $O(\varepsilon^{-(1+\kappa)/4})$ in $[0, b_2]$ and $O(1)$ in $[b_2, b]$. In order to guarantee resonance for positive values of x we therefore need to ensure that the two terms on the right-hand side of (6.14) do not cancel. As in Case III of § 6 we assume that we know the value of $\gamma(a) - \gamma(b)$, where $\gamma(x)$ is defined by (6.18). Then using the asymptotic approximations for the parabolic cylinder functions of large positive argument and Lemma 4.4, we calculate that the ratio of the first term to the second term is given by

$$(7.02) \qquad (-)^{s-1}\{1+o(1)\}\frac{\bar{m}}{\bar{l}}\left\{\frac{P(b)}{P(a)}\right\}^{(\kappa+1)/4} e^{\gamma(a)-\gamma(b)}.$$

Hence on using (5.14) we see that the requisite condition is given by

$$(7.03) \qquad l\, e^{Q(a)-\gamma(a)}\{-aA_0(a)\}^{1/2}\{P(a)\}^{\kappa/4} \neq (-)^s m\, e^{Q(b)-\gamma(b)}\{bA_0(b)\}^{1/2}\{P(b)\}^{\kappa/4}.$$

The interval $[a, 0]$ may be investigated in a similar manner, and our findings are as follows. Let c be the point of $(a, 0)$ for which

$$P(c) = P(b),$$

and restrict a_1 so that $c < a_1 < 0$. Then $\Lambda_3 e^{-P(x)/\varepsilon} W_3(\zeta)$ grows uniformly (and exponentially) in $[a_1, a_2]$ provided that (7.03) applies. And $\Lambda_4 e^{-P(x)/\varepsilon} W_4(\zeta)$ is uniformly $O(\varepsilon^{-(1+\kappa)/4})$ in $[a_2, 0]$ and $O(1)$ in $[a, a_2]$.

Therefore *when* $P(a) > P(b)$ *resonance takes place, provided that* (7.03) *is satisfied.* This is the case, for example, if either l or m is zero, or if l and $(-)^{s-1}m$ have the same sign.

When the condition (7.03) is violated some cancellation takes place on the right-hand side of (6.14) as $\varepsilon \to 0$; similarly for $\Lambda_3 e^{-P(x)/\varepsilon} W_3(\zeta)$. The present analysis is insufficiently delicate to determine whether resonance occurs in these circumstances however.

Case II. $P(b) > P(a)$. This case is treatable in a similar manner, and we find that *when* $P(b) > P(a)$ *resonance takes place, provided that* (7.03) *is satisfied.*

Case III. $P(a) = P(b)$. Equations (6.21) and (6.22) remain valid. Hence (6.25) applies, subject to the condition (6.24). Accordingly, $\Lambda_1 e^{-P(x)/\varepsilon} W_1(\zeta)$ grows uniformly (and exponentially) in $[b_2, b_1]$. Next, (6.26) applies, and from it we conclude that $\Lambda_2 e^{-P(x)/\varepsilon} W_2(\zeta)$ is uniformly $O(\varepsilon^{-(1+\kappa)/4})$ in $[0, b_2]$ and uniformly $O(1)$ in $[b_2, b]$. The analysis for the interval $[a, 0]$ is similar, and the general conclusion is the same as in Cases I and II, that is, *resonance takes place, provided that* (6.24) (or (7.03)) *is satisfied.*

An example in which the condition (6.24) is violated, but a definite conclusion can still be reached, is again supplied by the case in which the coefficients $A(\varepsilon, x)$ and

$B(\varepsilon, x)$ in (1.01) are even in x, and $a=-b$. Similar analysis to that given at the end of § 6 shows that resonance does *not* take place in these circumstances.

8. Exact eigenvalues. In the first parts of §§ 6 and 7 we proved that if $\beta \equiv \frac{1}{4}\pi(1+\kappa)$ is independent of ε and 2β is not an integer multiple of π, then resonance does not take place. We also showed that if β is permitted to depend on ε, then except for certain special boundary conditions resonance occurs whenever β satisfies equation (6.08), that is,

(8.01) $$\sin 2\beta+(\rho_1+\rho_3)\cos\beta+(\sigma_1+\sigma_3)\sin\beta+\rho_1\sigma_3+\rho_3\sigma_1=O(e^{-2\varpi/\varepsilon})$$

as $\varepsilon \to 0$, where ϖ is any constant exceeding both $|P(a)|$ and $|P(b)|$. As we noted in § 6, this equation has roots

$$\beta=\beta_s(\varepsilon), \qquad s=1,2,3,\cdots,$$

where

$$\beta_s(\varepsilon)=\tfrac{1}{2}s\pi+O\{\varepsilon \ln(1/\varepsilon)\}$$

as $\varepsilon \to 0$. The corresponding form of the original differential equation (1.01) is obtained by substituting

$$\kappa=(4/\pi)\beta_s(\varepsilon)-1$$

in (2.02). For each value of s there is a continuum of roots $\beta_s(\varepsilon)$; therefore since $\beta_s(\varepsilon)$ is not unique it is not an analytic function of ε.

In the form of the original problem that was investigated rigorously by Ackerberg and O'Malley, [1, § 5] and O'Malley [13, §§ 3 and 4], it was assumed that the coefficients $A(\varepsilon,x)$ and $B(\varepsilon,x)$ in (1.01) are holomorphic functions of ε and x and possess asymptotic expansions as $\varepsilon \to 0$ for x in some complex neighborhood of the interval $[a, b]$ and for ε in a sector $0<|\varepsilon|\leqq \varepsilon_0$, $|\mathrm{ph}\,\varepsilon|\leqq \theta_0$, where $\theta_0>0$. With these assumptions it was proved that necessary conditions for resonance are given by (1.04) or (1.05), that is (in both cases), $\beta=\frac{1}{2}s\pi$, where s is a positive integer. Correspondingly, $\kappa=2s-1$; compare (5.18). As we have already noted, these results are consistent with those of §§ 6 and 7.

In the first part of the present section we investigate circumstances in which (8.01) possesses *exact* roots $\beta=\frac{1}{2}s\pi$, where s is a positive integer. In other words, we seek sufficient conditions for resonance when the conditions (1.04) or (1.05) are fulfilled. Our assumptions will be the same as in § 2; thus no conditions of analyticity are imposed on the coefficients in (1.01).

On substituting $\beta=\frac{1}{2}s\pi$ in (8.01), we obtain either

(8.02) $$(-)^{s/2}(\rho_1+\rho_3)+\rho_1\sigma_3+\rho_3\sigma_1=O(e^{-2\varpi/\varepsilon}), \qquad s \text{ even};$$

or

(8.03) $$(-)^{(s-1)/2}(\sigma_1+\sigma_3)+\rho_1\sigma_3+\rho_3\sigma_1=O(e^{-2\varpi/\varepsilon}), \qquad s \text{ odd}.$$

One obvious situation in which these conditions are satisfied is given by

(8.04) $$\rho_1=\sigma_1=\rho_3=\sigma_3=0.$$

This happens, for example, when the function $\bar{\psi}(\varepsilon,\kappa,\zeta)$ in (5.02) is identically zero; in this event the original differential equation (1.01) is transformable into Weber's equation.

A more general situation in which (8.02), or (8.03), is satisfiable identically for all sufficiently small values of ε arises when the error terms ρ_1, σ_1, ρ_3, and σ_3 are independent of ε. Apart from the case (8.04) this may happen when equation (1.01) has a nontrivial solution that is independent of ε.[3]

Turning to cases in which ρ_1, σ_1, ρ_3, and σ_3 may vary with ε, we observe that (8.02) and (8.03) are satisfied when

$$\bar{\psi}(\varepsilon, 2s-1, \zeta) = O(e^{-2\varpi/\varepsilon}) \tag{8.05}$$

as $\varepsilon \to 0$, uniformly with respect to $\zeta \in [\bar{a}, \bar{b}]$. For in this case by taking the balancing function $\Omega(\zeta\sqrt{2/\varepsilon})$ in (5.06) to be $1+|\zeta|\sqrt{2/\varepsilon}$, we easily deduce from the definitions (5.19) and the error bounds (5.04) that both ρ_1 and σ_1 are $O\{\varepsilon \ln(1/\varepsilon)e^{-2\varpi/\varepsilon}\}$; compare (5.07). *A fortiori* ρ_1 and σ_1 are $O(e^{-2\varpi/\varepsilon})$. Similarly, both ρ_3 and σ_3 are $O(e^{-2\varpi/\varepsilon})$.

Equation (8.05) is therefore an appropriate condition of sufficiency to be adjoined to (1.04) or (1.05).[4] The only additional requirement concerns the boundary conditions, as specified in §§ 6 and 7. It will be observed that (8.05) is not unlike the condition given by Ackerberg and O'Malley on p. 292 of [1] pertaining to a coefficient $\sigma(\varepsilon)$ that appears in a certain transformation of the differential equation. The important difference is that we have explicit formulas for $\bar{\psi}(\varepsilon, 2s-1, \zeta)$, given in §§ 4 and 5, whereas only an existence theorem and asymptotic expansion are available for Ackerberg and O'Malley's $\sigma(\varepsilon)$.

We are now in a position to construct new examples of resonant systems, as follows:

$$\varepsilon y'' - xy' + \{x + s - 1 - \varepsilon + \varepsilon C(\varepsilon, x)\}y = 0, \tag{8.06}$$

and

$$\varepsilon y'' + xy' + \{x + s - \varepsilon + \varepsilon C(\varepsilon, x)\}y = 0. \tag{8.07}$$

In both cases the boundary conditions are again given by $y(a) = l$ and $y(b) = m$, with $l^2 + m^2 \neq 0$. Here s is a fixed positive integer, and $C(\varepsilon, x)$ is any continuous function of ε and x that satisfies

$$C(\varepsilon, x) = O(e^{-2\varpi/\varepsilon}) \tag{8.08}$$

as $\varepsilon \to 0$, uniformly with respect to $x \in [a, b]$, where ϖ is any constant exceeding both $\frac{1}{4}a^2$ and $\frac{1}{4}b^2$.

To establish resonance properties for (8.06) and (8.07), we write

$$P(x) = \mp\tfrac{1}{4}x^2, \tag{8.09}$$

where the upper sign applies in the case of (8.06), and the lower sign in the case of (8.07). On making the substitutions

$$y = e^{-P(x)/\varepsilon}W, \qquad x = 2\varepsilon + \zeta\sqrt{2}, \tag{8.10}$$

we obtain, in both cases,

$$d^2W/d\zeta^2 = \{\varepsilon^{-2}(\zeta^2 - 2s\varepsilon + \varepsilon) - 2C(\varepsilon, 2\varepsilon + \zeta\sqrt{2})\}W. \tag{8.11}$$

This differential equation has the form (5.02) with $-2C(\varepsilon, 2\varepsilon + \zeta\sqrt{2})$ in the role of

[3] When $A(\varepsilon, x)$ and $B(\varepsilon, x)$ are independent of ε, such nontrivial solutions are necessarily linear functions of x or constants. Some examples have been supplied by Kreiss and Parter [4, § 3].

[4] The condition (8.05) could be weakened somewhat with more detailed analysis.

$\bar{\psi}(\varepsilon, \kappa, \zeta)$, and $\kappa = 2s-1$. The boundary conditions are evidently given by

$$W(\bar{a}) = l\, e^{P(a)/\varepsilon}, \qquad W(\bar{b}) = m\, e^{P(b)/\varepsilon}, \tag{8.12}$$

where

$$\bar{a} = (a-2\varepsilon)/\sqrt{2}, \qquad \bar{b} = (b-2\varepsilon)/\sqrt{2}. \tag{8.13}$$

Comparison with (5.13) shows that l and m are now playing the roles of $\bar{l}$ and $\bar{m}$, respectively, and $\lambda = \mu = 0$. Lastly, from (8.08) it is clear that the condition (8.05) is satisfied.

On applying the results of § 6 to (8.06) we immediately conclude that in Case I, that is (in the present instance), $|a| > b$, resonance occurs if $m \neq 0$ but not if $m = 0$. Similarly in Case II, given by $|a| < b$, resonance occurs if $l \neq 0$, but not if $l = 0$. To deal with Case III, given by $a = -b$, we first compare the second of (8.10) with equations (6.18) and (8.09). This shows that $\gamma(x) = -x$. Hence (6.23) becomes

$$l \neq (-)^s m\, e^{2b}, \tag{8.14}$$

and resonance takes place in Case III whenever this condition is satisfied.

In the case of (8.07) we apply the results of § 7. We find that resonance occurs for all boundary conditions, except possibly those satisfying

$$la^s e^a = mb^s e^b; \tag{8.15}$$

compare (7.02), with $\bar{l}$ replaced by l, and $\bar{m}$ replaced by m.

Equations (8.06) and (8.07) appear to be the first established examples of resonance that satisfy the original conditions of Ackerberg and O'Malley and do not fall into either of the categories mentioned in the fourth and fifth paragraphs of this section. That is, neither (8.06) nor (8.07) is transformable into Weber's equation, nor is there a nontrivial solution that is independent of ε.

9. Matkowsky's criterion. Recently, Matkowsky [7] proposed an interesting formal criterion for resonance, applicable when the coefficients in the original differential equation are analytic functions of the parameter ε and independent variable x. His approach is to construct the outer solution of the boundary-value problem via the method of matched asymptotic expansions, and the criterion adopted for resonance is that all the coefficients in the outer expansion must be bounded throughout the interior of the interval of integration, including the turning point $x = 0$. Some applications have since been made to nonlinear problems [9]. Before we consider the validity of the criterion, it is instructive to discuss one of Matkowsky's examples in some detail.

Section 2 of [7] contains four examples, in three of which resonance is claimed in certain circumstances. Two of these examples are forms of Weber's equation and when corrected as in [8] the conclusions agree with the results of the present paper. The other example of resonance is given by

$$\varepsilon y'' - x(1+x^2)y' + (2+12\varepsilon)y = 0. \tag{9.01}$$

In the light of our observations in the preceding section, this differential equation appears to be of unusual interest, because it is not a form of Weber's equation. Unfortunately, however, the analysis in [7] has not been carried sufficiently far for this example, and the conclusions arrived at are incorrect. More complete analysis is as follows.

The outer expansion for the solution of (9.01) is given by

$$y=\sum_{j=0}^{\infty} w_j(x)\varepsilon^j,$$

where

$$(9.02)\qquad x(1+x^2)w_j'(x)-2w_j(x)=12w_{j-1}(x)+w_{j-1}''(x),\qquad j=0,1,2,\cdots,$$

with $w_{-1}(x)\equiv 0$; compare [7, eqs. (19) and (20)]. Integration of (9.02) yields

$$w_j(x)=\frac{x^2}{1+x^2}\int \frac{12w_{j-1}(x)+w_{j-1}''(x)}{x^3}\,dx.$$

From this relation it is easily seen that the coefficients have the form

$$w_j(x)=c_0q_j(x)+c_1q_{j-1}(x)+c_2q_{j-2}(x)+\cdots+c_jq_0(x),$$

where $c_0, c_1, c_2, \cdots$ are arbitrary constants of integration. The first three functions $q_j(x)$ may be verified to be

$$q_0(x)=\frac{x^2}{1+x^2},\qquad q_1(x)=-\frac{1}{1+x^2}-\frac{5x^2}{(1+x^2)^2}-\frac{2x^2}{(1+x^2)^3},$$

$$q_2(x)=\frac{12}{1+x^2}+\frac{42x^2}{(1+x^2)^2}+\frac{33x^2}{(1+x^2)^3}+\frac{30x^2}{(1+x^2)^4}+\frac{12x^2}{(1+x^2)^5}$$
$$+\frac{108x^2}{1+x^2}\ln x-\frac{54x^2}{1+x^2}\ln(1+x^2).$$

Each of the coefficients $w_0(x)$, $w_1(x)$, and $w_2(x)$ is therefore bounded as $x\to 0$, as required by Matkowsky's criterion. On proceeding to the next coefficient, we perceive that each term in $q_0(x)$, $q_1(x)$, and $q_2(x)$ that is analytic at $x=0$ makes a bounded contribution to $w_3(x)$. This takes care of all terms except the penultimate member of $q_2(x)$. For this term we find that

$$\frac{x^2}{1+x^2}\int\frac{1}{x^3}\left(12+\frac{d^2}{dx^2}\right)\left(\frac{x^2\ln x}{1+x^2}\right)dx=\chi(x)-\ln x,$$

where $\chi(x)$ denotes a function that is bounded at $x=0$. Hence $w_3(x)$ is unbounded at $x=0$. Accordingly, when Matkowsky's criterion is applied correctly to this example, it indicates that resonance does not take place. This is contrary to the statement on p. 88 of [7], but in agreement with our present results.

In order to show that Matkowsky's criterion can lead to a false conclusion when applied to the class of problems considered originally by Ackerberg and O'Malley, we consider the following (and final) example, given by

$$(9.03)\qquad \varepsilon y''-xy'+(x+1-\varepsilon+\varepsilon e^{-2\varpi/\varepsilon})y=0;\qquad y(a)=l,\quad y(b)=m,$$

where ϖ is a positive constant. It is easily seen that the coefficients in this equation satisfy all the conditions of the Uniform Reduction Theorem quoted in [1, p. 291], including those of analyticity with respect to ε and x.

The differential equation in (9.03) is the special form of the example (8.06) with $s=2$ and $C(\varepsilon,x)=e^{-2\varpi/\varepsilon}$. The results of § 8 therefore show that when

$$(9.04)\qquad \varpi>\max\left(\tfrac14 a^2,\tfrac14 b^2\right)$$

resonance occurs except for certain special values of l and m. This conclusion is consistent with Matkowsky's criterion, because the outer expansion is found to have the form

$$y = \sum_{j=0}^{\infty} (c_j x - 2c_{j-1})\, e^x \varepsilon^j,$$

in which c_0, c_1, c_2, $\cdots$ are arbitrary constants, and $c_{-1} = 0$. Thus all coefficients in the outer expansion are free from singularity everywhere in $[a, b]$.

Let us now suppose that instead of (9.04) the constant ϖ satisfies the condition

$$0 < \varpi < \min\left(\tfrac{1}{4}a^2, \tfrac{1}{4}b^2\right). \tag{9.05}$$

On making the substitutions (8.10) with

$$P(x) = -\tfrac{1}{4}x^2, \tag{9.06}$$

we transform (9.03) into

$$d^2W/d\zeta^2 = \{\varepsilon^{-2}(\zeta^2 - 3\varepsilon) - 2e^{-2\varpi/\varepsilon}\}W; \tag{9.07}$$

compare (8.11). This rearranges into the form (5.02) if we take

$$\kappa = 3 + 2\varepsilon e^{-2\varpi/\varepsilon}, \qquad \bar{\psi}(\varepsilon, \kappa, \zeta) = 0. \tag{9.08}$$

The analysis of § 5 then applies with each of the error terms $\eta_1(\zeta)$, $\eta_2(\zeta)$, $\eta_3(\zeta)$, and $\eta_4(\zeta)$ replaced by zero. Furthermore, because the boundary conditions for (9.07) are given by (8.12) and (8.13), we again see from (5.13) that l and m are playing the roles of $\bar{l}$ and $\bar{m}$, respectively, and $\lambda = \mu = 0$.

From (5.18) and (9.08) it follows that

$$\beta = \pi + \tfrac{1}{2}\pi\varepsilon e^{-2\varpi/\varepsilon}.$$

Hence for small ε

$$\cos 2\beta = 1 + O(\varepsilon^2 e^{-4\varpi/\varepsilon}), \qquad \sin 2\beta = \{1 + O(\varepsilon^2 e^{-4\varpi/\varepsilon})\}\pi\varepsilon e^{-2\varpi/\varepsilon}.$$

Next, from (5.33), (5.34), and (9.06), we see that

$$U(-\bar{a}\sqrt{2/\varepsilon}) = O\{\varepsilon^{-\kappa/2}\, e^{-a^2/(2\varepsilon)}\}\bar{U}(-\bar{a}\sqrt{2/\varepsilon}),$$

$$U(\bar{b}\sqrt{2/\varepsilon}) = O\{\varepsilon^{-\kappa/2}\, e^{-b^2/(2\varepsilon)}\}\bar{U}(\bar{b}\sqrt{2/\varepsilon}).$$

Substituting in (5.27) by means of these results and referring to the condition (9.05), we derive

$$D = -\{1 + o(1)\}\pi\varepsilon\, e^{-2\varpi/\varepsilon}\bar{U}(-\bar{a}\sqrt{2/\varepsilon})\bar{U}(\bar{b}\sqrt{2/\varepsilon}).$$

Similarly (5.23) reduces to

$$D\Lambda_1 = \{1 + o(1)\}m\, e^{-b^2/(4\varepsilon)}\bar{U}(-\bar{a}\sqrt{2/\varepsilon}) - l\, e^{-a^2/(4\varepsilon)}\bar{U}(\bar{b}\sqrt{2/\varepsilon}).$$

Dividing the last two equations and using the first of (5.03), we obtain

$$\Lambda_1\, e^{-P(x)/\varepsilon} W_1(\zeta) = -\{1 + o(1)\}\frac{m}{\pi\varepsilon}\, e^{2\varpi/\varepsilon}\, e^{(x^2 - b^2)/(4\varepsilon)}\frac{U(\zeta\sqrt{2/\varepsilon})}{\bar{U}(\bar{b}\sqrt{2/\varepsilon})}$$

$$+ \{1 + o(1)\}\frac{l}{\pi\varepsilon}\, e^{2\varpi/\varepsilon}\, e^{(x^2 - a^2)/(4\varepsilon)}\frac{U(\zeta\sqrt{2/\varepsilon})}{\bar{U}(-\bar{a}\sqrt{2/\varepsilon})}.$$

Then letting $\varepsilon \to 0$ and referring to (5.33), (5.34), and (5.35), we see that the right-hand side vanishes uniformly in $[0, b]$.

For the second term in (5.29) we find that

$$\Lambda_2 e^{-P(x)/\varepsilon} W_2(\zeta) = -\{1+o(1)\}\frac{l}{\pi\varepsilon} e^{2\varpi/\varepsilon} e^{(x^2-a^2)/(4\varepsilon)} \frac{U(\bar{b}\sqrt{2/\varepsilon})\bar{U}(\zeta\sqrt{2/\varepsilon})}{\bar{U}(\bar{b}\sqrt{2/\varepsilon})\bar{U}(-\bar{a}\sqrt{2/\varepsilon})}$$

$$+\{1+o(1)\} m e^{(x^2-b^2)/(4\varepsilon)} \frac{\bar{U}(\zeta\sqrt{2/\varepsilon})}{\bar{U}(\bar{b}\sqrt{2/\varepsilon})}.$$

As $\varepsilon \to 0$, the first term on the right-hand side vanishes uniformly in $[0, b]$, and the second term vanishes uniformly in $[0, b_1]$, where, as before, $0 < b_1 < b$. The corresponding results for the interval $[a, 0]$ may be arrived at in a similar way or by symmetry: we find that $\Lambda_3 e^{-P(x)/\varepsilon} W_3(\zeta)$ vanishes uniformly in $[a, 0]$, and $\Lambda_4 e^{-P(x)/\varepsilon} W_4(\zeta)$ vanishes uniformly in $[a_1, 0]$, where $a < a_1 < 0$.

The results just obtained show that the system (9.03) never resonates when the condition (9.05) is satisfied. On the other hand, Matkowsky's criterion does not distinguish between the cases (9.04) and (9.05), and would therefore lead to a false conclusion in the latter case. It is also interesting to note that the system (9.03) satisfies the infinite set of necessary conditions for resonance found by Lakin [5] and Cook and Eckhaus [2], irrespective of the value of the positive constant ϖ.

10. Summary and conclusions. In this paper we have studied the asymptotic behavior, for small ε, of the solutions of the boundary-value problem given by (1.01) and (1.02), with differentiability and other conditions on the coefficients $A(\varepsilon, x)$ and $B(\varepsilon, x)$ stated in the opening paragraph of § 2. By transforming the differential equation into a form resembling Weber's equation (§§ 3 and 4), we constructed four asymptotic solutions, complete with error bounds, two of which are uniformly valid in the interval $0 \leqq x \leqq b$, and the other two are uniformly valid in the interval $a \leqq x \leqq 0$ (§ 5). These solutions were then combined to satisfy the boundary conditions, and the asymptotic behavior of the resulting solution as $\varepsilon \to 0$ was examined in the cases $A(\varepsilon, x) < 0$ (§ 6) and $A(\varepsilon, x) > 0$ (§ 7). Particular attention was paid to the phenomenon of resonance, that is, an unusual lack of decay (§ 6) or an unusual growth (§ 7) of the solution as $\varepsilon \to 0$. We showed that in the general case resonance occurs essentially in the same manner as in the special case in which (1.01) is transformable into Weber's equation. That is, corresponding to each value of the positive integer s, there is an infinite set of values of the parameter κ appearing in the expression (2.02) for $B(\varepsilon, x)$ for which resonance takes place. The values of κ depend on ε, and are not known explicitly, except in the case of Weber's equation. Asymptotic estimates of κ are supplied by (6.11), in which $\hat{\varepsilon} = \varepsilon \ln(1/\varepsilon)$.

Next (§ 8), we considered the problem in the form proposed originally by Ackerberg and O'Malley [1], in which the coefficients $A(\varepsilon, x)$ and $B(\varepsilon, x)$ are assumed to be holomorphic functions of ε and x. Hitherto, only necessary conditions for resonance have been found in these circumstances. We completed these results by deriving a sufficient condition, and we illustrated this condition by two examples. In the concluding section (§ 9) we discussed a formal criterion for resonance that has been proposed recently, and showed by means of an example that this test is incomplete in its present form.

We have concentrated on the resonance phenomenon in the present paper, since this has been the least understood feature of the problem. The separation of the solution into boundary layers, which is of physical significance, has been neglected because this has been discussed fully by earlier writers. Nor did we consider details of the interesting oscillatory behavior of the solution in the neighborhood of $x = 0$. If desired, all of this information can be recovered from the uniform asymptotic approximations we have constructed for the solution.

Finally, we comment that we have restricted attention to a finite interval $[a, b]$, since this has been the assumption of earlier writers. However, the theory of differential equations with coalescing turning points that we have used is also applicable with unbounded values of the independent variable. In consequence, the corresponding boundary-value problem with $a = -\infty$ and $b = \infty$ is solvable by similar methods. Indeed, the analysis in this case is somewhat simpler, because *inter alīa* the values of κ that correspond to resonance are eigenvalues of the system and comprise a discrete set.

Acknowledgments. The author is indebted to Dr. R. E. O'Malley for illuminating discussions concerning the implications of the analyticity assumption on the coefficients in the differential equation, and also for providing several references. The author also thanks Dr. B. J. Matkowsky and the referees for several improvements in the presentation of the results.

REFERENCES

[1] R. C. ACKERBERG AND R. E. O'MALLEY, JR., *Boundary layer problems exhibiting resonance*, Studies in Appl. Math., 49 (1970), pp. 277–295.

[2] L. P. COOK AND W. ECKHAUS, *Resonance in a boundary value problem of singular perturbation type*, Ibid., 52 (1973), pp. 129–139.

[3] P. P. N. DE GROEN, *Elliptic singular perturbations of first-order operators with critical points*, Proc. Roy. Soc. Edinburgh Ser. A, 74 (1974/75), pp. 91–113.

[4] H. O. KREISS AND S. V. PARTER, *Remarks on singular perturbations with turning points*, this Journal, 5 (1974), pp. 230–251.

[5] W. D. LAKIN, *Boundary value problems with a turning point*, Studies in Appl. Math., 51 (1972), pp. 261–275.

[6] R. Y. LEE, *On uniform simplification of linear differential equation in a full neighborhood of a turning point*, J. Math. Anal. Appl., 27 (1969), pp. 501–510.

[7] B. J. MATKOWSKY, *On boundary layer problems exhibiting resonance*, SIAM Rev., 17 (1975), pp. 82–100.

[8] ———, *Errata: On boundary layer problems exhibiting resonance*, Ibid., 18 (1976), p. 112.

[9] B. J. MATKOWSKY AND W. L. SIEGMANN, *The flow between counter-rotating disks at high Reynolds number*, SIAM J. Appl. Math., 30 (1976), pp. 720–727.

[10] NATIONAL BUREAU OF STANDARDS, *Handbook of Mathematical Functions*, Appl. Math. Ser. No. 55, M. Abramowitz and I. A. Stegun, eds., U.S. Govt. Printing Office, Washington, DC, 1964.

[11] F. W. J. OLVER, *Asymptotics and Special Functions*, Academic Press, New York and London, 1974.

[12] ———, *Second-order linear differential equations with two turning points*, Philos. Trans. Roy. Soc. London Ser. A, 278 (1975), pp. 137–174.

[13] R. E. O'MALLEY, JR., *On boundary value problems for a singularly perturbed differential equation with a turning point*, this Journal, 1 (1970), pp. 479–490.

[14] ———, *Introduction to Singular Perturbations*, Academic Press, New York and London, 1974.

[15] C. E. PEARSON, *On a differential equation of boundary layer type*, J. Math. and Phys., 47 (1968), pp. 134–154.

[16] Y. SIBUYA, *Asymptotic solutions of a system of linear ordinary differential equations containing a parameter*, Funkcial. Ekvac., 4 (1962), pp. 83–113.

[17] A. M. WATTS, *A singular perturbation problem with a turning point*, Bull. Austral. Math. Soc., 5 (1971), pp. 61–73.

[18] E. ZAUDERER, *Boundary value problems for a second order differential equation with a turning point*, Studies in Appl. Math., 51 (1972), pp. 411–413.

[19] A. ZWAAN, *Intensitäten im* Ca-*Funkenspektrum*, Arch. Néerlandaises Sci. Exactes Natur. Ser. 3A, 12 (1929), pp. 1–76.

SIAM J. MATH. ANAL.
Vol. 9, No. 2, April 1978
Copyright © 1978 Society for Industrial and Applied Mathematics

A PERIODICITY THRESHOLD THEOREM FOR SOME NONLINEAR INTEGRAL EQUATIONS*

ROGER NUSSBAUM†

Abstract. We consider a parametrized family of nonlinear integral equations $F_\tau(x)=x$ which arise in the theory of epidemics. We study the equations by means of bifurcation theory and establish a best possible value τ_0 (a "periodicity threshold") such that for $\tau>\tau_0$ the equation has positive, periodic solutions. The value of τ_0 is determined by the spectral radius $r(L_\tau)$ of L_τ, where L_τ is an associated family of linear operators, so we also study properties of the map $\tau\to r(L_\tau)$.

Introduction. In a recent paper [2] K. Cooke and J. Kaplan suggested a model for the spread of a disease in which seasonality was incorporated (as opposed to [1]). If τ is a positive constant and $f(t,x)$ is a nonnegative, continuous function which is ω-periodic in t, they were led to consider

$$x(t)=\int_{t-\tau}^{t} f(s,x(s))\,ds. \tag{1}$$

Cooke and Kaplan then defined a number β (dependent on f) such that for each $\tau>\beta$ equation (1) has a positive solution of period ω. However, numerical studies in [2] for the special case $f(s,x)=(1+\frac{1}{2}\sin 2\pi s)g(x)$, where $g(x)=x(1-x)$ for $0\leqq x\leqq 1$ and $g(x)=0$ otherwise, suggested that positive periodic solutions of (1) exist for $\tau>1$, although the number β is 2 in this case.

In this paper we shall consider a more general class of integral equations, namely

$$x(t)=\int_{t-\tau}^{t} P(t-s,\tau)f(s,x(s))\,ds=(F_\tau x)(t) \tag{2}$$

We shall study (2) by means of a global bifurcation theorem and prove the existence of a positive number τ_0 such that (2) has a positive periodic solution for $\tau>\tau_0$ and will, in general, have none for $\tau<\tau_0$. In particular we shall prove that $\tau_0=1$ for the example above (see Theorem 3 and Remark 5). In fact our results are more precise than just showing existence of periodic solutions for $\tau>\tau_0$: see Theorem 3.

Typically in bifurcation problems it is assumed that the nonlinear operator $F_\tau(x)$ has a Fréchet derivative L_τ at 0 and that the map $\tau\to L_\tau$ is at least continuously Fréchet differentiable. In the context of the general theory, the novelty of our results is that the map $\tau\to L_\tau$ is not Fréchet differentiable. Nevertheless, the map $\tau\to r(L_\tau)=$ the spectral radius of L_τ may be very regular; part of this paper is devoted to proving this and studying $\tau\to r(L_\tau)$.

The organization of this paper is as follows: In the first section we prove a global bifurcation theorem which generalizes Theorem 1.3 of [11]. In the second section we apply the bifurcation theorem to (2). We associate to F_τ its Fréchet derivative L_τ; τ_0 is a value of τ such that $r(L_\tau)=1$. In the third section we study the problem of computing $r(L_\tau)$, and in the fourth section we consider the differentiability of the map $\tau\to r(L_\tau)$. Sections 2, 3 and 4 can be read independently.

1. A global bifurcation theorem. Let X denote a real Banach space. By a "cone" K in X we shall mean a closed, convex subset K of X such that $x\in K$ implies that $tx\in K$ for all $t\geqq 0$ and $x\in K-\{0\}$ implies $-x\notin K$. We shall say the cone is "total" if

* Received by the editors February 26, 1976, and in final revised form May 27, 1977.

† Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903. This work was supported in part by the National Science Foundation.

the closed linear span of K is all of X, and we shall always assume our cones are total. If x and y are elements of X and $y-x \in K$, we shall write $x \leqq y$. If $L: X \to X$ is a bounded linear operator such that $L(K) \subset K$, we shall say that L is "positive". If L_2 and L_1 are bounded linear maps of X to X, we shall write $L_1 \leqq L_2$ if $L_2 - L_1$ is positive. If L is a bounded linear operator we shall always write $r(L)$ = spectral radius of $L = \lim_{n\to\infty} \|L^n\|^{1/n}$.

If $J=(0, \infty)$ consider a map $F: K \times J \to K$ such that $F(0, \tau)=0$. Following Rabinowitz's approach [16], we shall study S, the closure in $X \times J$ of $\{(x, \tau): x \in K - \{0\}, \tau \in J, F(x, \tau)=x\}$. Motivated by the example of (1) (which we shall eventually study in the space X of continuous, real-valued functions of period ω) we assume

H1. $F: K \times J \to K$ is a continuous map which takes bounded sets into precompact sets and is such that $F(0, \tau)=0$ for all $\tau \in J$. If $F(x_k, \tau_k)=x_k$ for a sequence such that $x_k \neq 0$ and $\tau_k \to 0$, then it follows that $\lim_{k\to\infty} \|x_k\| = +\infty$.

The last part of H1 is a technical condition which will ensure that a set S_0 to be considered in Theorem 1 is unbounded. The condition can be omitted at the expense of complicating the statement of Theorem 1.

We also need to assume that $F_\tau(x)=F(x, \tau)$ has a linearization L_τ at 0, but the example of (1) again shows that L_τ cannot be assumed to be of a simple form like $L_\tau = \tau B$. Thus we suppose

H2. For each $\tau \in J$ there exists a compact, positive linear operator L_τ with the property that for any compact interval $[c, d] \subset J$, $\lim_{\|x\|\to 0} (\|x\|^{-1})(\|F(x, \tau)-L_\tau(x)\|)=0$ uniformly in $\tau \in [c, d]$. The map $\tau \to L_\tau$ is continuous with respect to the linear operator norm. Recall that a compact operator is one taking bounded sets to precompact sets.

Finally we shall need a technical condition on the spectral properties of the family L_τ.

H3. There exists a countable family $\Lambda \subset J$ with no finite accumulation point such that $x \neq L_\tau(x)$ for $x \in K - \{0\}$ and $\tau \notin \Lambda$.

Remark 1. If K is a total cone in a Banach X and L is a compact, positive linear operator with positive spectral radius r, then the Krein–Rutman theorem implies that there exists an $x \in K - \{0\}$ such that $Lx = rx$; see the appendix of [17] for details. In particular, H3 then implies that $\Lambda_1 = \{\tau \in J: r(L_\tau) = \text{spectral radius of } L_\tau = 1\}$ has no finite accumulation points, since $\Lambda_1 \subset \Lambda$.

Remark 2. Simple examples with positive matrices show that Λ is, in general, smaller than $\{\tau: 1 \text{ is an eigenvalue of } L_\tau\}$. This is one advantage of working in K.

If H2 and H3 hold, then a direct argument shows that for each $\tau \notin \Lambda$, there exists a positive number $\rho(\tau)$ such that $F(x, \tau) \neq x$ for $0 < \|x\| \leqq \rho(\tau)$ and $x \in K$. If $B_\rho = \{x \in K: \|x\| < \rho\}$ and if $i_K(F_\tau, B_{\rho(\tau)})$ denotes the fixed point index of F_τ on $B_{\rho(\tau)}$ (see [12] for a summary of the properties of the fixed point index and [13] for more details), then just as in [11, p. 328], we define $\Delta(\tau_0) = \lim_{\tau\to\tau_0^+} i_K(F_\tau, B_{\rho(\tau)}) - \lim_{\tau\to\tau_0^-} i_K(F_\tau, B_{\rho(\tau)})$. The homotopy property of the fixed point index implies that $\Delta(\tau_0)=0$ for $\tau_0 \notin \Lambda$.

PROPOSITION 1. *Assume that* H1, H2 *and* H3 *hold. If* $\tau_0 \in \Lambda$, *there exists* $\delta > 0$ *such that one of three possibilities hold*: (a) $r(L_\tau) > 1$ *for* $\tau_0 < \tau < \tau_0 + \delta$ *and* $r(L_\tau) < 1$ *for* $\tau_0 - \delta < \tau < \tau_0$, (b) $r(L_\tau) < 1$ *for* $\tau_0 < \tau < \tau_0 + \delta$ *and* $r(L_\tau) > 1$ *for* $\tau_0 - \delta < \tau < \tau_0$ *or* (c) $r(L_\tau) - 1$ *is of constant sign for* $0 < |\tau - \tau_0| < \delta$. *Furthermore, in case* (a) *we have* $\Delta(\tau_0) = -1$, *in case* (b) $\Delta(\tau_0) = 1$ *and in case* (c) $\Delta(\tau_0) = 0$.

Proof. By using the homotopy and additivity properties of the fixed point index together with H2 and H3, we find that $i_K(F_\tau, B_{\rho(\tau)}) = i_K(L_\tau, B_{\rho(\tau)})$. It follows from H3 and Remark 1 that there exists $\delta > 0$ such that $L_\tau(x) \neq x$ for $0 < |\tau - \tau_0| \leqq \delta$ and $x \in K - \{0\}$ and $r(L_\tau) \neq 1$ for $0 < |\tau - \tau_0| \leqq \delta$. The homotopy property shows that

$i_K(L_\tau, B_{\rho(\tau)})$ is constant for $\tau_0<\tau<\tau_0+\delta$ and constant for $\tau_0-\delta\leqq\tau<\tau_0$. To complete the proof it suffices to show that if L is a compact, positive linear operator such that $Lx\neq x$ for $x\in K-\{0\}$ and $B=\{x\in K: \|x\|<\rho\}$, then $i_K(L,B)=1$ if $r(L)<1$ and $i_K(L,B)=0$ if $r(L)>1$. If $r(L)<1$, consider the homotopy tL for $0\leqq t\leqq 1$. If $tLx=x$ for $\|x\|=\rho$ and $0\leqq t\leqq 1$, we find that $r(L)\geqq 1$, and we have a contradiction. Thus the homotopy property implies that $i_K(L,B)=i_K(L_0,B)=1$, where $L_0x=0$ for all $x\in K$.

If $r(L)=r>1$, use the Krein–Rutman theorem to select $x_0\in K-\{0\}$ such that $Lx_0=rx_0$. Consider the homotopy $L_s(x)=Lx+sx_0$ for $s\geqq 0$. Exactly the argument used in the final paragraph on p. 329 of [11] shows that $L_s(x)\neq x$ for $\|x\|=\rho$, $x\in K$ and $s\geqq 0$. It follows from Lemma 1.1 of [14] that $i_K(L,B)=0$. □

With the aid of Proposition 1, our next theorem is a direct consequence of Theorem 1.2 in [11].

THEOREM 1. *Suppose that* H1, H2, *and* H3 *hold*; *for some* $\tau_0\in\Lambda$ *suppose that* $r(L_\tau)>1$ *for* $\tau>\tau_0$ *and* τ *near* τ_0 *and* $r(L_\tau)<1$ *for* $\tau<\tau_0$ *and* τ *near* τ_0 *(or vice versa). Let S denote the closure in* $K\times J$ *of* $\{(x,\tau)\in K\times J: x\neq 0$ *and* $F(x,\tau)=x\}$ *and* S_0 *denote the maximal connected component of S which contains* $(0,\tau_0)$. *Then it follows that* S_0 *is nonempty and either* S_0 *is unbounded or* S_0 *contains a point* $(0,\tau_1)$ *with* $\tau_1\in\Lambda$ *and* $\tau_1\neq\tau_0$. *If* S_0 *is bounded and* Λ_0 *denotes the finite set of* $\tau\in\Lambda$ *such that* $(0,\tau)\in S_0$, *it follows that* $\sum_{\tau\in\Lambda_0}\Delta(\tau)=0$.

Theorem 1 is proved with the aid of the fixed point index, which can be viewed as generalizing the Leray–Schauder degree. The idea of using the Leray–Schauder degree in bifurcation theory can be found in Krasnosel'skii [9]. An important improvement of Krasnosel'skii's ideas has been given by Rabinowitz [16], whose work provides the model for the abstract bifurcation theorems in [11]. A less general version of the formula $\sum_{\tau\in\Lambda_0}\Delta(\tau)=0$ is implicit in Rabinowitz's work and has been remarked explicitly by E. N. Dancer [5] and J. Ize [7].

If one is only interested in the question of existence of a nonzero solution for a given $\tau\in J$, the following theorem suffices for our applications here.

THEOREM 2. *Let K be a total cone in a Banach space X and* $F: K\to K$ *a continuous, compact (nonlinear) map such that* $F(0)=0$. *Assume that there exists a bounded, compact, positive linear map L such that* $F(x)=L(x)+R(x)$, *where* $\lim_{\|x\|\to 0}[\|R(x)\|/\|x\|]=0$. *Suppose that* $r(L)>1$ *and* $Lx\neq x$ *for* $x\in K-\{0\}$. *Finally, assume that there exists a constant M such that* $x-tF(x)\neq 0$ *for* $\|x\|=M$ *and* $0\leqq t\leqq 1$. *Then* $F(x)=x$ *for some* $x\in K$ *with* $0<\|x\|<M$.

Proof. There exists a $\rho>0$ such that if $B_\rho=\{x\in K: \|x\|<\rho\}$, then F has no fixed points in $\bar{B}_\rho-\{0\}$ and $i_K(F,B_\rho)=i_K(L,B_\rho)$. The argument mentioned in the proof of Proposition 1 shows that $i_K(L,B_\rho)=0$. The homotopy $tF(x)$ for $0\leqq t\leqq 1$ shows that $i_K(F,B_M)=1$. If $U=\{x\in K: \rho<\|x\|<M\}$, it follows from the additivity property of the index that $i_K(F,U)=1$ and hence that F has a fixed point $x\in U$. □

We also shall need a condition which assures that a map $F: K\to K$ has no nonzero fixed point.

PROPOSITION 2. *Let K be a cone in a Banach space X and* $F: K\to K$ *a map. Assume that there exists a positive linear map* $L: K\to K$ *such that* $r(L)<1$ *and* $F(x)\leqq L(x)$ *for all* $x\in K$. *Then* 0 *is the only fixed point of F.*

Proof. An easy induction argument shows that

$$F^n(x)\leqq L^n(x) \tag{3}$$

for every positive integer n and every $x\in K$. If $F(x)=x$ for $x\in K-\{0\}$, equation (3) implies that $L^n(x)-x\in K$ for all n. Since $r(L)<1$, it follows that $L^n(x)\to 0$ and $-x\in K$. This is a contradiction. □

2. An integral equation from the theory of epidemics. We shall use Theorems 1 and 2 to study equation (2). Since we shall always be looking for solutions of period ω, henceforth we shall always denote by X the Banach space of continuous, real-valued functions $x(t)$ which are periodic of period ω in the sup norm; elements of X will be viewed as defined on all of R. Similarly, we shall denote by Y the square-integrable, real-valued functions $y(t)$ on $[0, \omega]$ in the L^2 norm; again these are extended to R by periodicity. We shall write $\tilde{X}$ and $\tilde{Y}$ for the complexifications of X and Y respectively.

We suppose about f that

H4. $f: R\times R^+\to R^+$ is a continuous function such that $f(s, 0)=0$ for all s and $f(s+\omega, x)=f(s, x)$ for all $(s, x)\in R\times R^+$. There exists a continuous, strictly positive, periodic function $a(s)$ of period ω such that $\lim_{x\to 0^+}[f(s, x)/x]=a(s)$ uniformly in s. Finally, we suppose $\lim_{x\to\infty}[f(s, x)/x]=0$, uniformly in s.

There is a great deal of freedom possible in choice of hypotheses on $P(u, \tau)$. For simplicity we restrict ourselves to the following assumptions:

H5. $P: R\times R^+\to R^+$ is a bounded, nonnegative, continuous function such that $P(u, \tau)>0$ for $u\in[0, \tau]$ and $P(u, \tau_1)\geqq P(u, \tau_2)$ whenever $\tau_1\geqq\tau_2$. Furthermore, if $a(s)$ is the function in H4, $\lim_{\tau\to\infty}(\inf_t\int_{t-\tau}^t P(t-s, \tau)a(s)\,ds)$ is strictly greater than one.

Let K_2 denote the cone of nonnegative functions in Y and K_1 the cone of nonnegative functions in X (X and Y as in the first paragraph of this section). Given $y\in K_2$, define a function $F: K_2\to K_1$ by $(F_\tau y)(t)=\int_{t-\tau}^t P(t-s, \tau)f(s, y(s))\,ds=z(t)$. It is a simple calculation, which we leave to the reader, to show that z is periodic of period ω.

Note that in the proof of Lemma 1 below, the same symbol is used for the norm in X and Y.

LEMMA 1. *Assume that* H4 *and* H5 *are satisfied and define* $F(y, \tau)=F_\tau(y)$. *Then* F *is a continuous map of* $K_2\times R^+$ *to* K_1 *and* F *takes bounded sets in* $K_2\times R^+$ *to precompact sets in* K_1. *If* H4 *is weakened by only assuming that* $f: R\times R^+\to R^+$ *is continuous and of period* ω *in its first variable, then* F *takes bounded sets in* $K_1\times R^+$ *to precompact sets in* K_1, *and* F *is continuous.*

Proof. If $x\in X$ or Y define $(Gx)(s)=f(s, x(s))$. If H4 holds, then there exist constants A and B such that $f(s, x)\leqq A+Bx$ and using continuity of f one can show that G defines a continuous map of K_2 to K_2 which takes bounded sets to bounded sets. If H4 does not hold, G defines a continuous map from K_1 to K_2 which takes bounded sets to bounded sets. Define a linear map $A_\tau: Y\to X$ by $(A_\tau z)(t)=\int_{t-\tau}^t P(t-s, \tau)z(s)\,ds$. Since $F_\tau=A_\tau G$, it suffices to show that $A(z, \tau)=A_\tau(z)$ defines a continuous map from $Y\times R^+$ to X which takes bounded sets to precompact sets.

Let S be a bounded set in $Y\times R^+$ and select a constant M such that $\sup_t(\int_{t-\tau}^t y^2(s)\,ds)\leqq M$ and $\tau\leqq M$ whenever $(y, \tau)\in S$. From H5 it follows that there is a constant C such that $P(u, \tau)\leqq C$ for all (u, τ). If $t\in R$, $\sigma>0$ and $\|z\|\leqq M$, it follows by the Cauchy–Schwarz inequality that for an integer M' with $M'\omega\geqq\sigma$ we have

$$\left|\int_{t-\sigma}^t P(t-s, \tau)z(s)\,ds\right|\leqq CM'\sigma^{1/2}\|z\|. \tag{4}$$

Inequality (4) shows that $A(S)$ is bounded in X. To prove $A(S)$ is an equicontinuous family, select an integer N such that $N\omega\geqq M$ and select $\delta>0$ such that $|P(u_1, \tau_1)-P(u_2, \tau_2)|<[\varepsilon/(2N)](M^{-1/2})$ whenever $|u_1-u_2|<\delta$ and $|\tau_1-\tau_2|<\delta$ and $\|u_j\|\leqq M$, $\|\tau_j\|\leqq M$. Also select δ small enough that $CM\delta^{1/2}<\varepsilon/4$. It then follows that if $(z, \tau)\in S$

and $0<t_2-t_1<\delta$, we have

$$\left|\int_{t_1-\tau}^{t_1} P(t_1-s,\tau)z(s)\,ds-\int_{t_2-\tau}^{t_2} P(t_2-s,\tau)z(s)\,ds\right|$$

$$\tag{5} \begin{aligned} &\leqq\left|\int_{t_1-\tau}^{t_2-\tau}[P(t_1-s,\tau)]z(s)\,ds\right|+\left|\int_{t_2-\tau}^{t_1}(P(t_1-s,\tau)-P(t_2-s,\tau))z(s)\,ds\right| \\ &\qquad+\left|\int_{t_1}^{t_2}P(t_2-s,\tau)z(s)\,ds\right|\leqq C\delta^{1/2}M+\frac{\varepsilon}{2}+C\delta^{1/2}M \\ &\qquad\leqq\varepsilon. \end{aligned}$$

It follows that $A(S)$ is precompact. The proof that A is continuous follows by estimates like (5), and we leave it to the reader. □

Results like Lemma 1 are familiar: see [9, Chap. 1]. Probably Lemma 1 exists in explicit form somewhere in the literature.

LEMMA 2. *Let $f: R\times R^+\to R^+$ be a map which satisfies* H4 *except for the assumption that $\lim_{x\to\infty}[f(s,x)/x]=0$ uniformly in s. Suppose that* H5 *is satisfied. Then it follows that $F: K_1\to K_1$ satisfies* H1. *If* H4 *is satisfied, F thought of as a map from K_2 to K_2 again satisfies* H1.

Proof. In either case Lemma 1 shows that F is a continuous, compact map from $K_j\times R^+$ to K_j, where j equals 1 or 2. It remains to show that if

$$\tag{6} F(x_k,\tau_k)=x_k$$

where $\tau_k\to 0$ and $x_k\in K_j-\{0\}$, then $\|x_k\|\to\infty$. Suppose first that f satisfies the weakened form of H4, and $x_k\in K_1$. Select t_k such that $|x_k(t_k)|=\|x_k\|>0$. Then if $P(u,\tau)\leqq C$ for all (u,τ), we obtain

$$\tag{7} \|x_k\|=\left|\int_{t_k-\tau_k}^{t_k}P(t_k-s,\tau_k)f(s,x_k(s))\,ds\right|\leqq C\tau_kM_k$$

where we define $M_k=\sup\{f(s,x): 0\leqq x\leqq\|x_k\|, s\in R\}$. If $\|x_k\|$ is bounded from some subsequence x_{k_j} (which we identify with x_k by relabeling), then (7) implies that $\|x_k\|\to 0$. However, the weakened form of H4 shows that there exists a positive constant C_1 such that

$$\tag{8} M_k\leqq C_1\|x_k\|$$

if $\|x_k\|$ is small enough. If $C_1C\tau_k<1$, inequalities (7) and (8) give a contradiction.

If H4 is satisfied, there exists a constant B (different from that in the proof of Lemma 1) such that $f(s,y)\leqq By$ for all $(s,y)\in R\times R^+$. If (x_k,τ_k) is as above with $x_k\in K_2-\{0\}$, the Cauchy–Schwarz inequality gives (for $\tau_k<\omega$)

$$\tag{9} |x_k(t)|\leqq BC\tau_k^{1/2}\|x_k\|.$$

Squaring both sides of (9) and integrating, we obtain a contradiction if $B^2C^2\tau_k\omega<1$. □

The proof of Lemma 2 actually shows that if H4 and H5 are satisfied and the constants B and C are as above that the equation $F(x,\tau)=x$ has no solution (x,τ) with $x\in K_1-\{0\}$ and $\tau\in R^+$ if $\tau<(BC)^{-1}$ and no solution (x,τ) with $x\in K_2-\{0\}$ if $\tau<(BC)^{-2}$.

We now assume that H4 and H5 hold and define a linear map L_τ by

$$\tag{10} (L_\tau y)(t)=\int_{t-\tau}^{t}P(t-s,\tau)a(s)y(s)\,ds.$$

Exactly the same argument used in Lemma 1 shows that L_τ defines a bounded compact linear map from Y to X such that $L_\tau(K_2) \subset K_1$. If $B(Y, X)$ denotes the Banach space of bounded linear operators from Y to X in the operator norm, then it is easy to show that the map $\tau \to L_\tau \in B(Y, X)$ is continuous. Since X is continuously imbedded in Y, the above remarks show that L_τ can be considered as a positive, compact linear operator from X to X or from Y to Y and that in either case $\tau \to L_\tau$ is continuous in the operator norm. Finally, it is not hard to show that if we work in X and $[c, d]$ denotes a bounded interval, then

$$\lim_{\|x\| \to 0} (\|x\|^{-1})(\|F(x, \tau) - L_\tau(x)\|) = 0$$

uniformly in $\tau \in [c, d]$ (note that this may not be true in Y). Thus we obtain the following lemma, whose detailed proof we leave to the reader.

LEMMA 3. *Assume* H4 *and* H5 *and consider F as a map from K_1 to K_1 with $F(x, \tau) \equiv F_\tau(x)$ and L_τ as a map from X to X. Then* H2 *is satisfied.*

At this point we need to recall a definition and a result from [10, Chap. 2].

DEFINITION 1. Let C be a cone in a Banach space Z and $L: Z \to Z$ a bounded linear operator such that $L(C) \subset C$. If $u_0 \in C - \{0\}$, L is called *u_0-positive* if for every $x \in C - \{0\}$, there exist an integer n and positive constants α and β (all dependent on x) such that

$$\alpha u_0 \leqq L^n(x) \leqq \beta u_0.$$

Our next lemma follows directly from the results in Chapter 2 of [10] and the Krein–Rutman theorem already mentioned in Remark 1. In the statement of Lemma 4, recall that a cone K in a Banach space X is called "reproducing" if $X = \{x - y : x, y \in K\}$. The cones K_1 and K_2 are clearly reproducing.

LEMMA 4 (see [10, Chap. 2]). *Let C be a reproducing cone in a Banach space Z and $L: Z \to Z$ a compact linear operator such that $L(C) \subset C$. Assume that L is u_0-positive for some $u_0 \in C - \{0\}$. Then if $r(L) = r$, we have $r > 0$, and if $Lx = \lambda x$ for $x \in K - \{0\}$, it follows that $\lambda = r$. Furthermore, $\{x \in Z: (rI - L)^n x = 0$ for some positive integer $n\}$ is one dimensional and contains a nonzero element of C.*

The linear maps we shall consider are all u_0-positive; specifically we have

LEMMA 5. *Assume that* H4 *and* H5 *hold, define u_0 to be the function which is identically one, and define L_τ by the formula* (10). *Then for $\tau > 0$, L_τ is u_0-positive as a map from K_1 to K_1 or as a map from K_2 to K_2.*

Proof. First take a fixed $y \in K_1 - \{0\}$. We have to show that there exist positive constants α and β and an integer n such that

$$\alpha u_0 \leqq L_\tau^n y \leqq \beta u_0.$$

If n and α have been found, it suffices to take $\beta = \|L_\tau^n\| \|y\|$. If $y(t_0) > 0$, then by continuity there exists an $\varepsilon > 0$ such that $y(t) > 0$ for $t \in [t_0, t_0 + \varepsilon]$. Since we assume $P(t - s, \tau)a(s)$ continuous and strictly positive for $t - \tau < s \leqq t$ and every t, it follows from the form of (9) that $y_1(t) = (L_\tau y)(t)$ is strictly positive for $t \in [t_0, t_0 + \varepsilon + \tau]$. Repeating this argument n times, we find that $(L_\tau^n y)(t) = y_n(t)$ is strictly positive for $t \in [t_0, t_0 + \varepsilon + n\tau]$. If n is selected so that $n\tau > \omega$, then by periodicity of y_n, y_n is everywhere positive and is continuous, so α exists.

The case of $L_\tau: K_2 \to K_2$ reduces to the previous case if we recall that L_τ actually is a continuous map from $K_2 \to K_1$ and if we show $y \in K_2 - \{0\}$ implies that $L_\tau(y) \in K_1 - \{0\}$. We leave the details to the reader. □

Lemmas 4 and 5 imply that if $\Lambda=\{\tau\in(0,\infty): L_\tau x=x \text{ for some } x\in K_1-\{0\}\}$, then $\Lambda=\{\tau: r(L_\tau)=1\}$. As we shall now show, the restrictions on $P(u,\tau)$ were made to ensure that $r(L_\tau)$ is a strictly increasing, continuous function of τ, so that Λ consists of at most one point. One can weaken the assumptions on P (and in fact we have omitted some interesting examples), but then the determination of Λ becomes more difficult.

LEMMA 6. *Assume that* H4 *and* H5 *hold and define* $L_\tau: X\to X$ *by the formula* (9). *Then it follows that the map* $\tau\to r(L_\tau)$ *is a strictly increasing, continuous function of* τ *for* $\tau>0$ *such that* $\lim_{\tau\to\infty} r(L_\tau)>1$ (*possibly* $+\infty$).

Proof. For $0<\tau<\sigma$, we have to show that $r(L_\tau)=r_1<r(L_\sigma)=r_2$. According to Lemma 4 there exists an $x_1\in K_1-\{0\}$ such that $L_\tau x_1=r_1x_1$ and $r_1>0$. Since $L_\tau^n(x_1)=r_1^nx_1$, Lemma 5 implies that there exist positive constants b and c such that

$$b\leqq x_1(t)\leqq c. \tag{11}$$

H4 and H5 imply that there exists a positive constant α such that

$$P(t-s,\sigma)a(s)\geqq\alpha \tag{12}$$

for $s\in[t-\tau, t-\tau+\Delta]$, where $\Delta=(\sigma-\tau)/2$ and the constant α is independent of t. If we recall that $P(u,\sigma)\geqq P(u,\tau)$ and apply (11) and (12) we obtain

$$\begin{aligned}(L_\sigma x_1)(t)&\geqq(L_\tau x_1)(t)+\int_{t-\sigma}^{t-\tau}P(t-s,\sigma)a(s)x_1(s)\,ds\\ &\geqq(r_1x_1)(t)+(b\alpha\Delta)\geqq(\rho x_1)(t)\end{aligned} \tag{13}$$

where $\rho=r_1-\alpha bc^{-1}\Delta$. It follows from (13) that

$$L_\sigma^n x_1\geqq\rho^n x_1. \tag{14}$$

If the spectral radius of L_σ were less than ρ, then (14) would imply that $-x_1\in K_1$, a contradiction. Thus we have $r(L_\sigma)\geqq\rho>r_1$.

The kind of argument used above also shows that given $\varepsilon>0$ and τ, x_1 and r_1 as above, there exists a $\delta>0$ such that for $\tau-\delta<\sigma\leqq\tau$ one has

$$L_\sigma x_1\geqq(r_1-\varepsilon)x_1. \tag{15}$$

It follows that $r(L_\sigma)\geqq r_1-\varepsilon$ for $\tau-\delta<\sigma\leqq\tau$.

It is true in general that for a continuous family L_σ of bounded linear operators

$$\limsup_{\sigma\to\tau} r(L_\sigma)\leqq r(L_\tau) \tag{16}$$

(see [6, Thm. 3.1]). On the other hand, the remarks above show that

$$\liminf_{\sigma\to\tau} r(L_\sigma)\geqq r(L_\tau). \tag{17}$$

The inequalities (16) and (17) yield the continuity of $r(L_\sigma)$.

The final condition in H5 ensures that

$$L_\tau e\geqq\lambda e \tag{18}$$

for τ large enough; here λ is a constant greater than one and e is the function identically one. It follows that $r(L_\tau)\geqq\lambda>1$ for τ large. □

We can now prove an existence theorem for positive periodic solutions of (2).

THEOREM 3. *Assume that* H4 *and* H5 *hold and let* $L_\tau: X\to X$ *be defined by* (10). *There exists a unique* $\tau_0>0$ *such that* $r(L_\tau)=$ *spectral radius of* L_τ *satisfies* $r(L_\tau)<1$, *for*

$\tau<\tau_0$, $r(L_{\tau_0})=1$ and $r(L_\tau)>1$ for $\tau>\tau_0$. If S denotes the closure in $K_1\times(0,\infty)$ of $\{(x,\tau): x\in K_1-\{0\}, \tau>0, F(x,\tau)=x\}$ and S_0 denotes the connected component of S which contains $(0,\tau_0)$, then S_0 is nonempty and unbounded. The set S contains no element of the form $(0,\tau)$ for $\tau\neq\tau_0$, and if J is any bounded subinterval of $(0,\infty)$, the set $S\cap(K_1\times J)$ is bounded. If $\tau>\tau_0$, there exists a nonzero solution $\phi\in K_1$ of the equation $F_\tau(\phi)=\phi$.

Proof. We must show that with F_τ and L_τ defined as in this section the assumptions H1, H2 and H3 are satisfied. Lemmas 2 and 3 show that H2 and H3 hold. Lemmas 4 and 5 show that the set Λ of H3 consists in our case of the unique point τ_0 such that $r(L_{\tau_0})=1$. Furthermore, Lemmas 1 and 5 show that $\Delta(\tau_0)=-1$. It follows from Theorem 1 that S_0 is unbounded.

If $F(x_k,\tau_k)=x_k$ for a sequence (x_k,τ_k) such that $\lim_{k\to\infty}(x_k,\tau_k)=(0,\tau)$ with $x_k\neq 0$, then if we define $u_k=x_k(\|x_k\|^{-1})$ and use H2 we find that

$$L_\tau(u_k)-u_k\to 0. \tag{19}$$

Since L is compact, we can assume by taking a subsequence that $u_k\to u$ and $L_\tau(u)=u$. This is a contradiction unless $\tau=\tau_0$.

We must show that if J is a bounded interval, there exists a constant M such that $\|\phi\|\leqq M$ if $(\phi,\tau)\in S$ and $\tau\in J$. Let A be a constant such that $\|L_\tau\|\leqq A$ for $\tau\in J$ and let ε be a number such that $\varepsilon A<1$. One can show (using H4) that there exists a constant M such that whenever $x\in K_1$ and $\|x\|\geqq M$ it follows that

$$f(s,x(s))\leqq\varepsilon\|x\|. \tag{20}$$

If $(\phi,\tau)\in S$ and $\tau\in J$, then (20) implies that one must have $\|\phi\|<M$, because if $\|\phi\|>M$ one would have

$$\begin{aligned}\|\phi\|=\|F_\tau(\phi)\|&\leqq\|L_\tau\|\,\|f(s,\phi(s))\|\\&\leqq\varepsilon A\|\phi\|,\end{aligned} \tag{21}$$

a contradiction.

Finally we note that $F_\tau(\phi)=\phi$ has a nonzero solution for $\tau>\tau_0$. For by the above remarks, S_0 contains elements (ϕ,τ) with τ arbitrarily large or τ arbitrarily close to τ_0. By connectedness of S_0, for each $\tau>\tau_0$, S_0 contains an element (ϕ,τ); since $(0,\tau)\notin S$, $\phi\neq 0$. □

Remark 3. In view of Theorem 3 it is reasonable to conjecture that under further assumptions on f S_0 may in fact be a curve $(\phi(s),\tau(s))$ with $(\phi(0),\tau(0))=(0,\tau_0)$ parametrized by some parameter s. We can prove this statement near $(0,\tau_0)$ by an argument in the spirit of the Crandall–Rabinowitz work in [3], although the precise conditions of Theorem 1.7 in [3] are not satisfied: our function $F(x,\tau)$ is not twice continuously Fréchet differentiable because the map $\tau\to L_\tau=F_x(0,\tau)$ is not C^1 as a map into the space of bounded linear maps with the norm topology.

The questions of parametrizing S_0 globally or of proving uniqueness of any positive solution of $F_\tau(\psi)=\psi$ for each $\tau>0$ (or for some subset of R^+) are interesting but apparently difficult. We have obtained some uniqueness results in [15], which was written after this paper. The results in [15] prove uniqueness of positive solutions of (1) for any $\tau>1$ for the example $f(s,x)=(1+\frac{1}{2}\sin 2\pi s)g(x)$ mentioned in the Introduction. For this example H. Smith [18] independently proved uniqueness, but only for the range $1<\tau\leqq 2$. However, for many functions $f(s,x)$, the structure of S_0 is still not well analyzed.

Remark 4. In general τ_0 in Theorem 3 is best-possible. Precisely, if $f(s, x) \leqq a(s)x$ for all $x \geqq 0$, then one can show that

$$F_\tau(x) \leqq L_\tau(x) \tag{22}$$

for any $\tau > 0$ and $x \in K$. Proposition 2 implies that F_τ has no nonzero fixed points in K for $\tau < \tau_0$.

Remark 5. If the function P is identically one, one can give simple crude estimates for τ_0 in Theorem 3. If $\tau = n\omega$ observe that L_τ has e as an eigenfunction with corresponding eigenvalue

$$r(L_\tau) = n\left(\int_0^\omega a(s)\, ds\right). \tag{23}$$

It follows that if n is the first integer such that

$$1 \leqq n\left(\int_0^\omega a(s)\, ds\right), \tag{24}$$

then $\tau_0 \in ((n-1)\omega, n\omega]$ with $\tau_0 = n\omega$ if equality holds in (24). Applying this remark to the example $a(s) = (1 + \frac{1}{2}\sin 2\pi s)$ in the Introduction shows that $\tau_0 = 1$ there.

3. Estimating the spectral radius of L_τ. We shall consider here the problem of calculating $r(L_\tau)$ to within a specified accuracy. Our approach will be to reduce the problem to that of computing $r(P_n L_\tau)$, where P_n is a finite dimensional projection. The basic difficulties are already apparent for

$$(L_\tau x)(t) = \int_{t-\tau}^{t} a(s)x(s)\, ds \tag{25}$$

where $x \in X$ and a and X are as before, and for simplicity we shall restrict ourselves to (25).

We begin by defining a sequence $\{P_n\}$ of finite dimensional linear projections on the space X of continuous, ω-periodic functions. Given an integer $n \geqq 1$ define $t_j = j\omega/n$ for $0 \leqq j \leqq n$. If $x \in X$, define $P_n x = y$ by $y(t_j) = x(t_j)$ for $0 \leqq j \leqq n$ and

$$y(t) = \frac{t_{j+1} - t}{t_{j+1} - t_j} y(t_j) + \frac{t - t_j}{t_{j+1} - t_j} y(t_{j+1}) \tag{26}$$

for $t_j \leqq t \leqq t_{j+1}$. Observe that P_n is a linear projection whose range is n dimensional and that $P_n(K) \subset K$.

LEMMA 7. *Let $a(s)$ be a continuous, positive periodic function and define $\alpha(\tau) = \inf_t \{\int_{t-\tau}^t a(s)\, ds\}$ and $\beta(\tau) = \sup_t \{\int_{t-\tau}^t a(s)\, ds\}$. If L_τ is defined by* (25) *for $\tau > 0$, then for $n > 1$ the spectral radii of L_τ and $P_n L_\tau$ satisfy*

$$\alpha(\tau) \leqq r(L_\tau) \leqq \beta(\tau), \qquad \alpha(\tau) \leqq r(P_n L_\tau) \leqq \beta(\tau).$$

Proof. If L is any positive linear operator on X, then

$$r(L) = \lim_{k\to\infty} \|L^k\|^{1/k} = \lim_{k\to\infty} \|L^k e\|^{1/k} \tag{27}$$

where e is the function identically one. One can see that

$$\alpha(\tau)e \leqq L_\tau(e) \leqq \beta(\tau)e$$

and since $P_n e = e$, this implies that

$$\alpha(\tau)e \leqq P_n L_\tau(e) \leqq \beta(\tau)e.$$

Using these estimates and iterating k times with the operators L_τ and $P_n L_\tau$ respectively yields

$$(28) \qquad \begin{aligned} (\alpha(\tau))^k e &\leqq L_\tau^k(e) \leqq (\beta(\tau))^k e, \\ (\alpha(\tau))^k e &\leqq (P_n L_\tau)^k(e) \leqq (\beta(\tau))^k e. \end{aligned}$$

By combining (27) and (28) we obtain the lemma. □

To state the next lemma we need a definition.

DEFINITION 2. If $M \geqq 1$ is a constant and K is the cone of nonnegative functions in X, define K_M by $K_M = \{x \in K : \max_t x(t) \leqq M \min_t x(t)\}$.

It is easy to show that K_M is a cone in X, and with a little more work one can show that K_M is total if $M > 1$.

LEMMA 8. *If the projections $\{P_n\}$ are defined as before, then $P_n(K_M) \subset K_M$ for each $n \geqq 1$ and $M \geqq 1$. If j is a positive integer and $j\omega \leqq \tau < (j+1)\omega$, then $L_\tau(K) \subset K_M$, where $M = (j+1)/j$.*

Proof. The projections P_n preserve the ordering induced by K, so if $x \in K$ and $\alpha e \leqq x \leqq \beta e$, it follows that $\alpha e = \alpha P_n e \leqq P_n x \leqq \beta e$. Applying this to the case $x \in K_M$, $\alpha = \min_t x(t)$ and $\beta = \max_t x(t)$, we obtain the first part of lemma.

To prove the second part, observe that for $x \in X$

$$(29) \qquad \begin{aligned} j\int_0^\omega a(s)x(s)\,ds = \int_{t-j\omega}^t a(s)x(s)\,ds &< \int_{t-\tau}^t a(s)x(s)\,ds \\ &\leqq (j+1)\int_0^\omega a(s)x(s)\,ds. \quad \Box \end{aligned}$$

We have already seen that L_τ has a unique positive eigenvector of norm one with corresponding eigenvalue $r(L_\tau)$. The next proposition shows that the same thing is true for $P_n L_\tau$ if $\omega/n < \tau$ (if $\omega/n > \tau$, there may be nonzero vectors $x \in K$ such that $(P_n L_\tau)(x) = 0$).

PROPOSITION. *Let a be a positive, continuous periodic function of period ω and suppose L_τ is defined by (25). If $\omega/n < \tau$, there exists a unique $x_n \in K - \{0\}$ such that $\|x_n\| = 1$ and $P_n L_\tau x_n = \lambda_n x_n$; λ_n is necessarily $r(P_n L_\tau)$.*

Proof. By the remarks in the previous section, it suffices to show $P_n L_\tau$ is e-bounded. To prove e-boundedness, it suffices to show that if $x \in K - \{0\}$, then $(P_n L_\tau)^k(x)$ is strictly positive everywhere for some integer k. Since $x \in K - \{0\}$, select s such that $x(s) > 0$. The form of $L_\tau x = y$ shows that y is strictly positive on $[s, s+\tau]$, and since $\tau > \omega/n$, there must exist a point $t_j = j\omega/n \in [s, s+\tau]$. This shows that $P_n y$ is positive at t_j. Assume inductively that we have shown $(P_n L_\tau)^m(x)$ is strictly positive on some interval $[t_j, t_{j+m-1}]$. Then the same reasoning as above shows $(P_n L_\tau)^{m+1}$ is strictly positive on $[t_j, t_{j+m}]$, so that $(P_n L_\tau)^{n+1}(x)$ is everywhere strictly positive. □

Because of the above proposition, given a fixed $\tau > 0$ and n so large that $\omega/n < \tau$ we can define x_n to be the unique element of K of norm one such that $P_n L_\tau x_n = \lambda_n x_n$ ($\lambda_n = r(P_n L_\tau)$) and x to be the unique element of K of norm one such that $L_\tau x = \lambda x$. Of course x and x_n depend on τ.

We shall see that it is crucial for our problem to estimate the minimum of x_n and x, so we make the following definition.

DEFINITION 3. If x_n and x are as above, define $k_n(\tau) = (\min_t x_n(t))^{-1}$ and $k(\tau) = (\min_t x(t))^{-1}$.

LEMMA 9. *Let $a(s)$ be as in Lemma 7 and let $\alpha(\tau)$ and $\beta(\tau)$ be defined as in Lemma 7. If $j\omega \leqq \tau < (j+1)\omega$ for a positive integer j, then $k(\tau) \leqq (j+1)/j$ and $k_n(\tau) \leqq (j+1)/j$ for $\omega/n < \tau$. If $0 < \tau < \omega$, it follows that $k(\tau) \leqq \exp[\|a\|\omega/(\alpha(\tau))]$ and if also $\|a\|\omega/(n\alpha(\tau)) < \frac{1}{2}$, then $k_n(\tau) \leqq \exp[2\|a\|\omega/(\alpha(\tau))]$.*

Proof. If $j\omega \leqq \tau < (j+1)\omega$ and $M = (j+1)/j$, then Lemma 8 implies that $x_n = \lambda_n^{-1} P_n L_\tau x_n \in K_M$ and $x = \lambda^{-1} L_\tau x \in K_M$, so the first part of the lemma is immediate.

If $0 < \tau < \omega$, we have that

$$x(t) = \lambda^{-1} \int_{t-\tau}^{t} a(s)x(s)\, ds$$

where $\lambda = r(L_\tau)$ and $\alpha(\tau) \leqq \lambda$. Differentiating the above equation and using the fact that $x(s) > 0$ for all s gives

$$x'(t) = \lambda^{-1}[a(t)x(t) - a(t-\tau)x(t-\tau)] \leqq \frac{\|a\|}{\alpha(\tau)} x(t). \tag{30}$$

Suppose x has its minimum at t_0. Then the equation

$$\frac{d}{dt}\left[x(t)\left(\exp\left(-\frac{\|a\|}{\alpha(\tau)}\right)\left(t - t_0\right)\right)\right] \leqq 0$$

implies that for $t_0 \leqq t$

$$x(t) \leqq x(t_0) \exp\left(\frac{\|a\|}{\alpha(\tau)}(t - t_0)\right) \tag{31}$$

Since x is periodic of period ω, equation (31) gives the estimate on $k(\tau)$.

The argument for $k_n(\tau)$ follows the same outline, but with some technical complications. Assume that $\|a\|\omega/(n\alpha(\tau)) < \frac{1}{2}$ (which implies $\omega/n < \tau$) and for positive integers j define $t_j = j\omega/n$. We know that x_n is linear on each interval $[t_j, t_{j+1}]$ and that

$$x_n'(t) = \left(\frac{\omega}{n}\right)^{-1} [x_n(t_{j+1}) - x_n(t_j)] \quad \text{for } t_j \leqq t \leqq t_{j+1}.$$

The definitions of x_n, $\lambda_n = r(P_n L_\tau)$ and P_n yield

$$x_n(t_{j+1}) = \lambda_n^{-1} \int_{t_{j+1}-\tau}^{t_{j+1}} a(s)x_n(s)\, ds,$$

$$x_n(t_j) = \lambda_n^{-1} \int_{t_j-\tau}^{t_j} a(s)x_n(s)\, ds.$$

These equations give that

$$\begin{aligned} x_n(t_{j+1}) - x_n(t_j) &= \lambda_n^{-1}\left[\int_{t_j}^{t_{j+1}} a(s)x_n(s)\, ds - \int_{t_j-\tau}^{t_{j+1}-\tau} a(s)x_n(s)\, ds\right] \\ &\leqq \lambda_n^{-1}\left[\int_{t_j}^{t_{j+1}} a(s)x_n(s)\, ds\right]. \end{aligned}$$

It follows from the above inequality that

$$(\omega/n)^{-1}[x_n(t_{j+1}) - x_n(t_j)] \leqq \lambda_n^{-1}\|a\| \max(x_n(t_j), x_n(t_{j+1})).$$

If $x_n(t_j) \leqq x_n(t_{j+1})$, the previous inequality implies

$$x_n(t_{j+1}) \leqq \left(\frac{\omega}{n}\right)\left(\frac{\|a\|}{\lambda_n}\right) x_n(t_{j+1}) + x_n(t_j)$$

and since we are assuming $\|a\|\omega/(n\alpha(\tau))<\frac{1}{2}$ we obtain

$$x_n(t_{j+1})\leqq 2x_n(t_j),$$

which is clearly also true if $x_n(t_j)>x_n(t_{j+1})$. Thus we find that if $x_n(t_j)\leqq x_n(t_{j+1})$ and $t_j\leqq t\leqq t_{j+1}$ we have

$$x_n'(t)\leqq\frac{\|a\|}{\lambda_n}x_n(t_{j+1})\leqq\frac{2\|a\|}{\lambda_n}x_n(t) \tag{32}$$

and this inequality is trivially valid if $x_n(t_j)>x_n(t_{j+1})$.

Now suppose that x_n has its minimum at t_k. By using inequality (32) (recalling that $\lambda_n\geqq\alpha(\tau)$), we easily obtain as before that

$$x_n(t)\leqq x_n(t_k)\exp\left(\frac{2\|a\|(t-t_k)}{\alpha(\tau)}\right)$$

and using the periodicity of x_n, we obtain from this the estimate on $k_n(\tau)$. □

Remark 6. The estimate in Lemma 9 for $k(\tau)$ approaches infinity as $\tau\to 0$. If x_τ denotes the unique normalized positive eigenvector of L_τ, $x_\tau=e$ if a is a constant function and $k(\tau)=1$ in that case. If a is not a constant function, one can show that there does not exist a convergent sequence (in X) of eigenvectors x_{τ_n} with $\tau_n\to 0$. This suggests $k(\tau)$ may behave badly as $\tau\to 0$, but we have no sharp results on the behavior of $k(\tau)$ for small τ. However, we are primarily interested in finding τ such that $r(L_\tau)=1$, and for these purposes we can restrict attention to τ such that $\beta(\tau)\geqq 1$. For this range of τ, Lemma 9 provides a *computable* constant k such that $k(\tau)\leqq k$ and $k_n(\tau)\leqq k$.

Our next lemma is a simple application of the mean value theorem which we leave to the reader.

LEMMA 10. *Suppose that $x\in X$, x is differentiable and $|x'(t)-x'(s)|\leqq C|t-s|$ for all t, s. Then we have*

$$\|P_nx-x\|<(C/2)(\omega/n)^2.$$

We are now in a position to prove an approximation theorem for $r(L_\tau)$.

THEOREM 4. *Let a be a positive periodic function and assume that a is Lipschitzian with Lipschitz constant C. Define $\alpha(\tau)$ and $\beta(\tau)$ as in the statement of Lemma 7 and $k(\tau)$ and $k_n(\tau)$ as in Definition* 3. (Recall that Lemma 9 gives estimates for $k(\tau)$ and $k_n(\tau)$). *Then it follows that*

$$r(L_\tau)\geqq r(P_nL_\tau)-\left[C+\frac{\|a\|^2}{\alpha(\tau)}\right]\left(\frac{\omega}{n}\right)^2k_n(\tau),$$

$$r(L_\tau)\leqq r(P_nL_\tau)+\left[C+\frac{\|a\|^2}{\alpha(\tau)}\right]\left(\frac{\omega}{n}\right)^2k(\tau).$$

Proof. Suppose $\lambda x=L_\tau x$ for $x\in K$, $\|x\|=1$. If $y=L_\tau x$, then y is differentiable everywhere and

$$y'(t)=a(t)x(t)-a(t-\tau)x(t-\tau).$$

It follows from the above equation that

$$\begin{aligned}|y'(t)-y'(s)|<{}&|a(t)-a(s)|\,|x(t)|+|a(s)|\,|x(t)-x(s)|\\&+|a(t-\tau)-a(s-\tau)|\,|x(t-\tau)|\\&+|a(s-\tau)|\,|x(t-\tau)-x(s-\tau)|.\end{aligned} \tag{33}$$

Since $|x'(t)|=(1/\lambda)|y'(t)|\leqq\|a\|/(\alpha(\tau))$, inequality (33) implies that

$$|y'(t)-y'(s)|\leqq 2\left[C+\frac{\|a\|^2}{\alpha(\tau)}\right]|t-s|=2C_1|t-s|. \tag{34}$$

Lemma 10 and inequality (34) imply that

$$\|P_n y-y\|\leqq C_1(\omega/n)^2=\varepsilon_n. \tag{35}$$

Inequality (35) and the definition of $k(\tau)$ now give that

$$P_nL_\tau x\geqq\lambda x-\varepsilon_n\geqq(\lambda-\varepsilon_n k(\tau))x. \tag{36}$$

Since P_nL_τ is a positive operator, the inequality (36) implies that

$$r(P_nL_\tau)\geqq r(L_\tau)-\varepsilon_n k(\tau), \tag{37}$$

which is the second inequality in the statement of Theorem 4.

To prove the remaining inequality, let x_n denote the normalized eigenvector of P_nL_τ in K and write $y_n=L_\tau x_n$. If one recalls that the argument used in Lemma 9 shows that

$$|x_n'(t)|\leqq\frac{\|a\|}{\alpha(\tau)}\quad\text{for } t\in\left[\frac{j\omega}{n},\frac{(j+1)\omega}{n}\right],$$

then essentially the same argument used above shows

$$|y_n'(t)-y_n'(s)|\leqq 2\left[C+\frac{\|a\|^2}{\alpha(\tau)}\right]|t-s|.$$

Lemma 10 now implies that (writing $\lambda_n=r(P_nL_\tau)$)

$$\|P_ny_n-y_n\|=\|\lambda_nx_n-L_\tau x_n\|\leqq\left[C+\frac{\|a\|^2}{\alpha(\tau)}\right]\left(\frac{\omega}{n}\right)^2=\delta_n.$$

This implies that

$$L_\tau x_n\geqq\lambda_nx_n-\delta_n\geqq[\lambda_n-\delta_nk_n(\tau)]x_n,$$

and since L_τ is a positive linear operator we conclude that

$$r(L_\tau)\geqq r(P_nL_\tau)-\left[C+\frac{\|a\|^2}{\alpha(\tau)}\right]\left(\frac{\omega}{n}\right)^2k_n(\tau). \quad\square$$

Remark 7. To compute the eigenvalues of P_nL_τ, let $X_n=P_n(X)=$ an n dimensional space and find the eigenvalues of $P_nL_\tau|X_n$, an $n\times n$ matrix. This reduces the question to an algebraic one. However, the matrix for P_nL_τ will in general have no nonzero entries even for a simple function like $a(s)=c+d\sin 2\pi s$, so the method may be unsuitable for large n. By exploiting the fact that L_τ: $K\to K_M$ for $\tau\geqq\omega$, it is possible (for $\tau\geqq\omega$) to give approximation theorems of the type $|r(Q_nL_\tau^j)-r(L_\tau^j)|\leqq M\varepsilon_n$, where $\{Q_n\}$ is a sequence of finite dimensional linear projections such that $\|Q_nL_\tau^j-L_\tau^j\|=\varepsilon_n\to 0$ and j is a fixed positive integer. If a is Lipschitzian, one can take, for example, j to be 2 and Q_n to be the orthogonal projection (considered as a map of X into X) onto the span of $\cos kx$, $\sin kx$, $0\leqq k\leqq n$. The matrix $Q_nL_\tau^j$ may be simpler than P_nL_τ. For reasons of length we omit details.

4. Differentiability of the map $\tau \to r(L_\tau)$. Consider a linear map $L_\tau: X \to X$ (a map more general than in § 2) defined by

$$(38) \qquad (L_\tau x)(t) = \int_{t-\tau}^{t-g(\tau)} P(t-s, \tau)a(s)x(s)\,ds.$$

We assume about the functions a, P and g in (38)

H6. $a: \mathbb{R} \to \mathbb{R}$ is a $C^n (n \geqq 0)$, strictly positive periodic function of period ω. The map $P: \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$ is such that $(\partial/\partial s)^j (\partial/\partial \tau)^k P(s, \tau)$ exists and is continuous for $j+k \leqq n+1$, and P is bounded, nonnegative and strictly positive for $s \in (g(\tau), \tau)$. The function $g: \mathbb{R}^+ \to \mathbb{R}^+$ is C^{n+1} and $0 \leqq g(\tau) < \tau$ for $\tau > 0$.

We shall prove that under assumption H6 the map $\tau \to r(L_\tau)$ is C^{n+1}. Our interest in this question arises from several sources. For instance, consider a mild generalization of equations in § 2:

$$x(t) = \int_{t-\tau}^{t-\tau/2} f(s, x(s))\,ds.$$

To apply Theorem 1 to this equation, we have to know that if

$$\Lambda = \{\tau: r(L_\tau) = 1\}, \quad \text{where } (L_\tau x)(t) = \int_{t-\tau}^{t-\tau/2} a(s)x(s)\,ds,$$

then Λ has no finite accumulation points. However, it is unclear whether $r(L_\tau)$ is a monotonic increasing function of τ in this situation, and investigating the structure of Λ is much harder than in § 2. Information about derivatives of $r(L_\tau)$ helps in studying the structure of Λ. Of course if one could establish real analyticity of $r(L_\tau)$ for a, g and P real analytic, Λ would necessarily have no accumulation points, but calculations for the simple case $g \equiv 0$, $P \equiv 1$ and $a(s) = 1 + b \sin \pi s$ with $0 < |b| < 1$ suggest that real analyticity may not hold, although we have no proof yet.

The question of differentiability also arises in a more general model for epidemics suggested by H. Smith [18]. Smith is led to the linear map

$$(L_{\alpha,\beta} x)(t) = \int_{t-\beta}^{t-\alpha} a(s)P(t-s)x(s)\,ds$$

where $0 < \alpha < \beta$, a is as in H6 and $P(u) \geqq \delta > 0$ for all $u \geqq 0$. For each $\alpha > 0$ one defines $\beta(\alpha) = \beta$ to be the unique β such that $r(L_{\alpha,\beta}) = 1$, and one is interested in properties of the map $\beta(\alpha)$. The techniques we develop here can be used to show $\alpha \to \beta(\alpha)$ is C^{n+1} if a and P are C^n and to study other properties of $\beta(\alpha)$, although we do not pursue this here.

The proof we shall give of differentiability is rather long, and before embarking on it we should mention the technical difficulties which arise even for $P \equiv 1$ and $g \equiv 0$. In this case the map $\tau \to L_\tau$ is not Fréchet differentiable (in the norm topology for operators) even if $a(s)$ is a positive constant. If $a(s)$ is not a constant, the operators L_τ and L_σ do not in general commute. Finally, the operator L_τ is not normal (for nonconstant functions a) when considered on the Hilbert space $L^2(a(s)ds)$ of square-integrable ω-periodic functions x with norm

$$\left(\int_0^\omega x^2(s)a(s)\,ds\right)^{1/2},$$

which seems the natural Hilbert space in which to work. The adjoint L_τ^* of L_τ is

$$(L_\tau^* x)(t) = \int_t^{t+\tau} a(s)x(s)\,ds,$$

which need not commute with L_τ.

We now begin the proof. The same sort of argument used § 2 shows that L_τ is e-bounded (e is the function identically one) if H6 is satisfied. It follows from Lemma 4 that if $\lambda = r(L_\tau)$, then $\bigcup_{j\geqq 1} N((\lambda - L_\tau)^j)$ = the generalized null space of $\lambda - L_\tau$ is one dimensional. If $\tilde{X}$ denotes the complexification of X, L_τ is defined in the natural way on $\tilde{X}$ and if Γ is a simple closed curve in the complex plane which contains no elements of the spectrum of L_τ, one can define a linear projection Q_τ by

$$Q_\tau = \frac{1}{2\pi i}\int_\Gamma (\zeta - L_\tau)^{-1}\,d\zeta. \tag{39}$$

If D denotes the interior of Γ and if $\lambda = r(L_\tau)$ is the only point of the spectrum of L_τ in $\bar{D}$, it is known (see [8, pp. 178–181]) that $R(Q_\tau)$; the range of Q_τ, is the generalized null space of $\lambda - L_\tau$. Using this fact and Theorem 6.17 (on p. 178 of [8]) one also sees that if W is the range of $\lambda - L_\tau$ in $\tilde{X}$ and $S = \lambda - L_\tau | W$, then W has codimension one and S is one-to-one and onto.

Our next lemma will be convenient in studying the projections Q_τ; the result is known but we include a proof for completeness.

LEMMA 11. *Let P_1 and P_2 be bounded linear projections on a Banach space Z and denote their ranges by R_1 and R_2. Then if $\|P_1 - P_2\| < 1$, $\dim R_1 = \dim R_2$ (if one is infinite dimensional, both are).*

Proof. $A = I + (P_2 - P_1)$ and $B = I + (P_1 - P_2)$ are one-to-one maps of X onto X, because $\|P_2 - P_1\| < 1$. Since $A(R_1) \subset R_2$ and $B(R_2) \subset R_1$, it follows that $\dim R_1 \leqq \dim R_2$ and $\dim R_2 \leqq \dim R_1$. □

LEMMA 12. *Assume* H6, *let σ be a fixed positive number and let $\lambda = r(L_\sigma)$ with corresponding positive eigenvector x_0 or norm one. Define W to be the range of $\lambda - L_\sigma$ in $\tilde{X}$ and $S^{-1} = \lambda - L_\sigma | W$ (as a map of W onto W). Let f be a continuous linear functional such that $f(x_0) = 1$ and $f|W = 0$ and define $Q(x) = x - f(x)x_0$. Let Γ denote a circle of radius $r = (\|f\|^{1/2})(\|Q\|^{1/2} + \|f\|^{1/2})^{-1}(\|S\|)^{-1}$ about λ in $\mathbb{C}$. Then $\zeta - L_\tau$ is one-to-one, onto X if $\zeta \in \Gamma$ and*

$$\|L_\tau - L_\sigma\| < (\|S\|)^{-1}(\|Q\|^{1/2} + \|f\|^{1/2})^{-2} = C^{-1}. \tag{40}$$

Further, if J is an interval of reals containing σ such that (40) *holds for $\tau \in J$ and if Q_τ is defined by* (39), *then Q_τ has one dimensional range for $\tau \in J$.*

Proof. Suppose that $(\zeta - L_\sigma)x = y$ for $\zeta \in \Gamma$. We can write $x = ax_0 + u$ and $y = bx_0 + v$, where u, $v \in W$. The equation becomes

$$a(\zeta - \lambda)x_0 + S^{-1}[I + (\zeta - \lambda)S]u = bx_0 + v. \tag{41}$$

Solving (41) gives

$$x = \frac{b}{\zeta - \lambda}x_0 + [I + (\zeta - \lambda)S]^{-1}Sv \tag{42}$$

and since

$$[I + (\zeta - \lambda)S]^{-1} \leqq \frac{1}{1 - |\zeta - \lambda|\,\|S\|}, \quad bx_0 = f(y) \quad \text{and} \quad v = Q(y),$$

we obtain

$$\|x\| = \|(\zeta - L_\sigma)^{-1} y\| \leqq \left[\frac{\|f\|}{r} + \frac{\|S\|\,\|Q\|}{1 - r\|S\|}\right]\|y\|. \tag{43}$$

Simplifying the bracketed expression on the right side of (43) gives

$$\|(\zeta - L_\sigma)^{-1}\| \leqq \|S\|(\|Q\|^{1/2} + \|f\|^{1/2})^2. \tag{44}$$

If we write

$$\zeta - L_\tau = [I - (L_\tau - L_\sigma)(\zeta - L_\sigma)^{-1}](\zeta - L_\sigma)$$

and assume that (40) holds, then $\zeta - L_\tau$ is a product of invertible operators and so invertible.

To prove $\dim R(Q_\tau) = 1$ for $\tau \in J$, note that $\tau \to Q_\tau$ is continuous in the norm topology. Thus if for each nonnegative integer n (including $n = +\infty$) we define $E_n = \{\tau \in J : \dim R(Q_\tau) = n\}$, Lemma 11 implies E_n is open (as a subset of J). Thus E_1 is nonempty ($\sigma \in E_1$), open and closed (it is the complement in J of $\bigcup_{n \neq 1} E_n$). Since J is connected, $E_1 = J$. □

Now let notation and assumptions be as in the statement and proof of Lemma 12 and let J_1 be an open interval containing σ such that (40) holds and such that

$$\|L_\tau - L_\sigma\| < \varepsilon \tag{45}$$

for $\tau \in J_1$, where ε will be selected later. According to Lemma 12, $R(Q_\tau)$ is one dimensional, and since $L_\tau Q_\tau = Q_\tau L_\tau$, it follows that if x_τ is a nonzero element of $R(Q_\tau)$, x_τ is an eigenvector of L_τ with eigenvalue λ_τ. One can easily show (we omit the proof) that $Q_\tau(X) \subset X$, so that if $Q_\tau(x_0) \neq 0$, then $Q_\tau(x_0) \in X$ is an eigenvector of L_τ with real eigenvalue λ_τ. However, if inequality (45) holds we have

$$\begin{aligned} Q_\tau &= \frac{1}{2\pi i}\int_\Gamma (\zeta - L_\sigma)^{-1}[I - (L_\tau - L_\sigma)(\zeta - L_\sigma)^{-1}]^{-1}\, d\zeta \\ &= \frac{1}{2\pi i}\sum_{k=0}^{\infty}\int_\Gamma (\zeta - L_\sigma)^{-1}[(L_\tau - L_\sigma)(\zeta - L_\sigma)^{-1}]^k\, d\zeta. \end{aligned} \tag{46}$$

If C is as in inequality (40) and $\delta > 0$ is such that $\delta < \|f\|^{-1}$, then (46) implies that if $(\varepsilon C^2)(1 - \varepsilon C)^{-1} < 1$ we have

$$\|Q_\tau - Q_\sigma\| < \delta. \tag{47}$$

Since we have

$$f(Q_\tau(x_0)) > f(Q_\sigma x_0) - \delta\|f\|\,\|x_0\| = 1 - \delta\|f\| > 0 \tag{48}$$

it follows that $Q_\tau(x_0) \neq 0$ and that

$$\lambda_\tau = [f(L_\tau Q_\tau(x_0))][f(Q_\tau x_0)]^{-1}. \tag{49}$$

We claim that $Q_\tau(x_0)$ is an element of the interior of K for all $\tau \in J_1$. This is certainly true for τ near σ, since $x_0 \in \mathring{K}$. To prove it in general suppose not, and let $\tau_1 \in J_1$ be the first $\tau \in J_1$ such that $\tau > \sigma$ and $Q_\tau(x_0) \notin \mathring{K}$. By our remarks $Q_{\tau_1}(x_0) = z \neq 0$, $z \in K$ and z is an eigenvector of L_{τ_1}. However, since L_{τ_1} is e-positive, we must have $z \in \mathring{K}$, a contradiction.

Since $Q_\tau(x_0) \in \mathring{K}$ for $\tau \in J_1$ and since it is an eigenvector of L_τ, it follows that $\lambda_\tau = r(L_\tau)$. Thus to prove that $r(L_\tau)$ is a C^{n+1} function of τ, it suffices to show that $Q_\tau(x_0)$ and $L_\tau Q_\tau(x_0)$ are C^{n+1} functions of τ. Equality (46) provides an expression for

$Q_\tau(x_0)$; to study $L_\tau Q_\tau(x_0)$ observe that for $\tau \in J_1$ we have

$$\begin{aligned} L_\tau Q_\tau &= -I + \frac{1}{2\pi i} \int_\Gamma \zeta(\zeta - L_\tau)^{-1}\, d\zeta \\ &= -I + \frac{1}{2\pi i} \sum_{k=0}^{\infty} \int_\Gamma \zeta(\zeta - L_\sigma)^{-1}[(L_\tau - L_\sigma)(\zeta - L_\sigma)^{-1}]^k\, d\zeta. \end{aligned} \tag{50}$$

We need some further notation. For simplicity we shall henceforth write X instead of $\tilde{X}$. For each nonnegative integer j let X_j denote the Banach space of j times continuously differentiable, complex-valued periodic functions of period ω and set $X_{-1} = X_0 = X$; we define the norm on X_j by

$$\|x\|_j = \sum_{m=0}^{j} \sup_t |x^{(m)}(t)|.$$

If $L\colon X_j \to X_k$ is a bounded linear operator, we shall write $\|L\|_{j,k}$ to denote the norm of L as an element of $B(X_j, X_k)$, the space of bounded linear operators from X_j to X_k.

Our next lemma follows essentially by a direct (though messy) differentiation.

LEMMA 13. *Let notation be as above and suppose that* $g\colon \mathbb{R}^+ \to \mathbb{R}^+$ *is* C^{n+1}, $a\colon \mathbb{R} \to \mathbb{R}$ *is* C^n *and periodic of period* ω *and* $P\colon \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$ *is such that* $(\partial/\partial s)^j(\partial/\partial \tau)^k P(s, \tau)$ *exists and is continuous for* $j + k \leq n + 1$ (weaker assumptions than in H6). *If* L_τ *is defined by* (38), *it follows that* $L_\tau \in B(X_j, X_{j+1})$ *for* $0 \leq j \leq n$ *and* $L_\tau \in B(X_j, X_j)$ *for* $j = n + 1$. *If* $u \in X_j$ *for* $j \leq n$ *and* $0 \leq m \leq j + 1$, *then the map* $\tau \to L_\tau u \in X_{j+1-m}$ *has* m *derivatives in the* X_{j+1-m} *topology and* $(d^{(m)}/d\tau^{(m)})(L_\tau u) = L_\tau^{(m)}(u)$, *where* $L_\tau^{(m)} \in B(X_j, X_{j+1-m})$ (we omit reference to j in the notation $L_\tau^{(m)}$). *Furthermore the map* $\tau \to L_\tau^{(m)}$ *is continuous in the strong operator topology on* $B(X_j, X_{j+1-m})$.

Proof. We induct on the integer n in the statement of the lemma. If $n = 0$ the result is clear once one observes that for $u \in X_0$,

$$\begin{aligned} \left(\frac{d}{d\tau} L_\tau u\right)(t) &= \int_{t-\tau}^{t-g(\tau)} \frac{\partial}{\partial \tau} P(t-s, \tau) a(s) u(s)\, ds \\ &\quad + P(\tau, \tau) a(t-\tau) u(t - \tau - g'(\tau)) - P(g(\tau), \tau) a(t - g(\tau)) u(t - g(\tau)) \end{aligned} \tag{51}$$

where the derivative in (51) is taken in the X_0 topology, and

$$\begin{aligned} \frac{d}{dt}((L_\tau u)(t)) &= \int_{t-\tau}^{t-g(\tau)} \frac{\partial}{\partial t} P(t-s, \tau) a(s) u(s)\, ds \\ &\quad + P(g(\tau), \tau) a(t - g(\tau)) u(t - g(\tau)) - P(\tau, \tau) a(t-\tau) u(t-\tau). \end{aligned} \tag{52}$$

Generally, assume inductively the lemma is true for any functions a, g and P which satisfy the hypotheses of the lemma for $n < N$ and suppose that a, g and P satisfy the differentiability assumptions for $n = N$. If $u \in X_N$, equation (52) holds. Clearly the last two terms on the right in (52) possess N derivatives in t which are bounded by $K_1\|u\|_N$, K_1 independent of u. By inductive hypothesis (thinking of $(\partial/\partial t)P(t-s, \tau)$ as $P_1(t-s, \tau)$) the integral in (52) determines a continuous linear map from X_N to X_N and hence has N derivatives in t bounded by $K_2\|u\|_N$. Thus we have $L_\tau \in B(X_N, X_{N+1})$, and since X_{N+1} is continuously imbedded in X_N, $L_\tau \in B(X_{N+1}, X_{N+1})$. The same inductive argument shows $L_\tau \in B(X_j, X_{j+1})$ for $0 \leq j < N$.

The second part of the lemma follows by a similar inductive argument once one observes that for $j \leq N$ (51) holds for $u \in X_j$ if $a \in C^N$, $g \in C^{N+1}$ and $P \in C^{N+1}$ (the derivative with respect to τ being taken in the X_j topology). We leave details to the reader. □

Since X_{j+1} is compactly imbedded in X_j for $j \geqq 0$, Lemma 13 implies that $L_\sigma: X_j \to X_j$ is a compact linear operator for $j \leqq n$. If L is a compact linear map of a Banach space Z into itself and ζ is a nonzero scalar, $\zeta - L$ is one-to-one and onto iff it is one-to-one. Since $\zeta - L$ is one-to-one and onto X_0 for $\zeta \in \Gamma$, it follows that $\zeta - L_\sigma$ is one-to-one and onto as a map of X_j for $0 \leqq j \leqq n$.

Before studying the differentiability of $Q_\tau x_0$ and $L_\tau Q_\tau x_0$, we need some lemmas concerning the differentiability of the terms in the series expansion (46) of $Q_\tau x_0$.

LEMMA 14. *Let U, V, W be Banach spaces with $U \subset V \subset W$ and continuous inclusion maps. Suppose that J is an interval of reals, $u: J \to U$ is a continuous map and $\tau \to i(u(\tau)) \in V$ is a C^1 map, where $i: U \to V$ is inclusion. Assume that we are given a map $\tau \to A_\tau \in B(V, W)$ which is continuous in the strong operator topology on $B(V, W)$ and that for each fixed $u_0 \in U$, the map $\tau \to A_\tau(i(u_0)) \in W$ is C^1. Then $\tau \to A_\tau(i(u(\tau)))$ is a C^1 map into W. Further, if we write $(d/d\tau)(A_\tau(i(u_0))) = A_\tau^{(1)}(u_0)$ and $u^{(1)}(\tau) = (d/d\tau)(iu(\tau))$, we have*

$$\frac{d}{d\tau} A_\tau(iu(\tau)) = A_\tau(u^{(1)}(\tau)) + A_\tau^{(1)}(u(\tau)).$$

Proof. Since no confusion should result, we shall omit the inclusion i in our formulas. To prove the lemma it suffices to show the following equalities hold (where limits are taken with respect to the W topology):

$$\lim_{\Delta \to 0} A_{\tau+\Delta}\left(\frac{u(\tau+\Delta) - u(\tau)}{\Delta}\right) = A_\tau(u^{(1)}(\tau)),$$

$$\lim_{\Delta \to 0} \left(\frac{A_{\tau+\Delta} - A_\tau}{\Delta}\right)(u(\tau)) = A_\tau^{(1)}(u(\tau)).$$

The second equality holds by assumption. To prove the first, observe that the continuity of A_τ in the strong operator topology and the uniform boundedness principle imply that there exists a constant M such that $\|A_{\tau+\Delta}\| \leqq M$ and

$$\left\| A_{\tau+\Delta}\left(\frac{u(\tau+\Delta) - u(\tau)}{\Delta} - u^{(1)}(\tau)\right)\right\|_W \leqq M \left\| \frac{u(\tau+\Delta) - u(\tau)}{\Delta} - u^{(1)}(\tau)\right\|_V \to 0.$$

The strong operator continuity also gives that

$$\lim_{\Delta \to 0} A_{\tau+\Delta}(u^{(1)}(\tau)) = A_\tau(u^{(1)}(\tau)),$$

so the lemma is proved. □

In our next lemma we adhere to the notation of Lemma 13; also for notational convenience we set $S_\tau = L_\tau - L_\sigma$ and $S_\tau^{(m)} = L_\tau^{(m)}$ for $1 \leqq m \leqq n+1$.

LEMMA 15. *Assume that* H6 *holds and that n is as in* H6. *Suppose that J is an interval of reals and we are given a map $v: J \to X_n$ such that v has j continuous derivatives in the X_{n-j} topology for $0 \leqq j \leqq n+1$* (recall that $X_{-1} = X_0$). *For $\zeta \in \Gamma$ define $w(\tau) \in X_n$ by*

$$w(\tau) = (\zeta - L_\sigma)^{-1}(L_\tau - L_\sigma)(v(\tau)).$$

Then $w(\tau)$ has the same differentiability properties as $v(\tau)$ and

$$w^{(j)}(\tau) = (\zeta - L_\sigma)^{-1} \sum_{k=0}^{j} \binom{j}{k} S_\tau^{(j-k)}(v^{(k)}(\tau)). \tag{53}$$

Proof. In the notation of Lemma 14, set $U = X_n$, $V = W = X_{n-1}$. The map $\tau \to v(\tau) \in X_{n-1}$ is C^1 and $(\zeta - L_\sigma)^{-1} S_\tau$: $V \to W$ satisfies the conditions of Lemma 14, so $w(\tau)$ is a C^1 map into X_{n-1} and

$$w^{(1)}(\tau) = (\zeta - L_\sigma)^{-1}[S_\tau^{(1)}(v(\tau)) + S_\tau(v^{(1)}(\tau))].$$

We assume inductively that we have shown that $\tau \to w(\tau) \in X_{n-m}$ has m derivatives for $1 \leqq m \leqq j \leqq n$, with the derivatives given by the formula (53), and we try to prove the same result for $j+1$. By assumption we have

$$w^{(j)}(\tau) = (\zeta - L_\sigma)^{-1} \sum_{k=0}^{j} \binom{j}{k} S_\tau^{(j-k)}(v^{(k)}(\tau)).$$

For a fixed $k \leqq j$, define $U = X_{n-k}$, $V = X_{n-k-1}$ and $W = X_{n-j-1}$ and observe that the conditions of Lemma 14 hold for the term $S_\tau^{(j-k)}(v^{(k)}(\tau))$. Thus we can differentiate term by term and observe that

$$\begin{aligned} w^{(j+1)}(\tau) &= (\zeta - L_\sigma)^{-1} \sum_{k=0}^{j} \binom{j}{k} S_\tau^{(j-k+1)}(v^{(k)}(\tau)) + (\zeta - L_\sigma)^{-1} \sum_{k=0}^{j} \binom{j}{k} S_\tau^{(j-k)}(v^{(k+1)}(\tau)) \\ &= (\zeta - L_\sigma)^{-1} \sum_{k=0}^{j+1} \binom{j+1}{k} S_\tau^{(j+1-k)}(v^{(k)}(\tau)), \end{aligned}$$

which is of the required form. □

We are finally in a position to prove our theorem about the differentiability of the map $\tau \to r(L_\tau)$.

THEOREM 5. *Assume that* H6 *holds, let X denote the space of continuous, complex-valued functions of period ω and for $\tau > 0$ define a map L_τ: $X \to X$ by the formula* (38). *Then if n is as in* H6, *the map $\tau \to r(L_\tau) =$ the spectral radius of L_τ is $n+1$ times continuously differentiable.*

Proof. By our previous remarks, it suffices to show that the maps $\tau \to Q_\tau(x_0)$ and $\tau \to L_\tau Q_\tau(x_0)$ are C^{n+1}. Let $J \subset J_1$ be an open interval containing σ and for $\tau \in J$ and $\zeta \in \Gamma$ define

$$w_k(\tau, \zeta) = [(\zeta - L_\sigma)^{-1} S_\tau]^k ((\zeta - L_\sigma)^{-1}(x_0)).$$

It follows from Lemma 13 that an eigenvector of L_σ which corresponds to a nonzero eigenvalue must lie in X_{n+1}, so $x_0 \in X_{n+1}$. Lemma 15 and formula (53) now imply that $\tau \to w_k(\tau, \zeta)$ has j continuous derivatives in the X_{n-j} topology for $0 \leqq j \leqq n+1$ and that

$$\frac{d^j}{d\tau^j} w_k(\tau, \zeta) = w_k^{(j)}(\tau, \zeta)$$

is a continuous function of ζ in the X_{n-j} topology. Notice that

$$(54)\quad Q_\tau(x_0) = \frac{1}{2\pi i} \sum_{k=0}^{\infty} \int_\Gamma w_k(\tau, \zeta)\, d\zeta \quad \text{and} \quad L_\tau Q_\tau(x_0) = x_0 + \frac{1}{2\pi i} \sum_{k=0}^{\infty} \int_\Gamma \zeta w_k(\tau, \zeta)\, d\zeta$$

where the integrals are taken in X_0. Suppose we can prove that for $0 \leqq j \leqq n+1$

$$(55)\qquad \|w_k^{(j)}(\tau, \zeta)\|_{X_{n-j}} \leqq \varepsilon_k$$

where $\sum_{k=0}^{\infty} \varepsilon_k < \infty$ and ε_k is independent of $\zeta \in \Gamma$ and $\tau \in J$. Then a standard argument shows that $Q_\tau(x_0)$ and $L_\tau Q_\tau(x_0)$ are C^{n+1} in X_0 with derivatives obtained by differentiating the series in (54) term by term.

Thus it remains to prove an estimate like (55). Recall that the subscript on a norm indicates the space in which the norm is taken. By formula (53) we have for $0 \leqq j \leqq n+1$

$$\begin{aligned} \|w_k^{(j)}(\tau,\zeta)\|_{n-j} \leqq & \sum_{m=0}^{j-1} \binom{j}{m} \|(\zeta - L_\sigma)^{-1} S_\tau^{(j-m)}\|_{(n-m,n-j)} \|w_{k-1}^{(m)}(\tau,\zeta)\|_{n-m} \\ & + \|(\zeta - L_\sigma)^{-1} S_\tau\|_{n-j} \|w_{k-1}^{(j)}(\tau,\zeta)\|_{n-j}\,. \end{aligned} \tag{56}$$

Take a constant K such that for all $\zeta \in \Gamma$, $\tau \in J$ and integers m and j with $0 \leqq m < j \leqq n+1$ we have

$$\|(\zeta - L_\sigma)^{-1} S_\tau^{(j-m)}\|_{(n-m,n-j)} \leqq K.$$

By taking the interval J about σ to be sufficiently small we can assume that for all $\zeta \in \Gamma$, $\tau \in J$ and integers j with $0 \leqq j \leqq n+1$ we have

$$\|(\zeta - L_\sigma)^{-1} S_\tau\|_{n-j} \leqq c < 1.$$

If we define $\mu_k^{(j)}$ for $0 \leqq j \leqq n+1$ and $k \geqq 0$ by the formula

$$\mu_k^{(j)} = \sup \{\|w_k^{(j)}(\tau,\zeta)\|_{n-j} : \tau \in J,\ \zeta \in \Gamma\}$$

then (55) implies

$$\mu_k^{(j)} \leqq \sum_{m=0}^{j-1} \binom{j}{m} K \mu_{k-1}^{(m)} + c\mu_{k-1}^{(j)}. \tag{57}$$

Let c_1 be such that $c < c_1 < 1$ and suppose $0 \leqq j \leqq n+1$. We claim that there exists a constant A such that

$$\mu_k^{(j)} \leqq k(k-1)\cdots(k-j+1)Ac_1^{k-j} \tag{58}$$

for $k \geqq n+1$; when $j=0$, the right hand side of (58) is interpreted as Ac_1^k. By our previous remarks we will be done if we can prove (58). Select an integer $k_1 \geqq n+1$ such that for any integer $k \geqq k_1$ and for $0 \leqq j \leqq n+1$ we have

$$\sum_{m=0}^{j-1} \binom{j}{m} K \frac{(k-1)(k-2)\cdots(k-m)}{k(k-1)\cdots(k-j+1)} < c_1 - c. \tag{59}$$

This can be done because each term in the summation is dominated by $\binom{j}{m}Kk_1^{-1}$. Next select A so large that formula (58) holds for $0 \leqq j \leqq n+1$ and for $n+1 \leqq k \leqq k_1$. We assume inductively that formula (58) is valid for $k \geqq k_1$ and for $0 \leqq j \leqq n+1$ and try to prove it for $k+1$; formula (57) gives

$$\begin{aligned} \mu_{k+1}^{(j)} \leqq & \sum_{m=0}^{j-1} \binom{j}{m} K(k)\cdots(k-m+1)Ac_1^{k-m} \\ & + (cAc_1^{k-j})(k)(k-1)\cdots(k-j+1) \\ \leqq & (Ac_1^{k-j})(k+1)\cdots(k-j+2)\left\{c + \sum_{m=0}^{j-1}\binom{j}{m} K \frac{(k)\cdots(k-m+1)}{(k+1)(k)\cdots(k-j+2)}\right\} \\ \leqq & (Ac_1^{k-j})(k+1)\cdots(k-j+2)\{c_1\}. \end{aligned} \tag{60}$$

In the above formulas $(k)\cdots(k-m+1)$ is interpreted as 1 for $m=0$. Inequality (60) completes the inductive proof. □

Remark 8. If one is only interested in proving $\tau \to r(L_\tau)$ is C^1 if a is continuous and P is C^1 (the case $n=0$), a simpler proof using the implicit function theorem and following the idea of Crandall and Rabinowitz in [4] can be given. This approach fails to give higher order differentiability in our case, however.

Remark 9. The argument we have given here can be abstracted to the situation of a parametrized family of linear operators L_τ which operate on a nested family of Banach spaces $X_{n+1} \subset X_n \subset \cdots \subset X_0$ and which are assumed to satisfy the conclusions of Lemma 14 and some other conditions. The explicit form of (38) has not been strongly used.

REFERENCES

[1] K. L. COOKE AND J. A. YORKE, *Equations modelling population growth and gonorrhea epidemiology*, Math. Biosci., 31 (1976), pp. 87–105.

[2] K. L. COOKE AND J. L. KAPLAN, *A periodicity threshold theorem for epidemics and population growth*, Math. Biosci., 31 (1976), pp. 87–105.

[3] M. G. CRANDALL AND P. H. RABINOWITZ, *Bifurcation from simple eigenvalues*, J. Functional Anal., 8 (1971), 321–340.

[4] ———, *Bifurcation, perturbation of simple eigenvalues, and linearized stability*, Arch. Rational Mech. Anal., 52 (1973), pp. 161–180.

[5] E. N. DANCER, *On the structure of solutions of nonlinear eigenvalue problems*, Indiana Univ. Math. J. 23 (1974), pp. 1069–1076.

[6] F. HOPPENSTEADT AND P. WALTMAN, *A problem in the theory of epidemics II*, Math. Biosci., 12 (1971), pp. 133–145.

[7] J. IZE, *Bifurcation theory for Fredholm operators*, Mem. Amer. Math. Soc., 174 (1976).

[8] T. KATO, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1966.

[9] M. A. KRASNOSEL'SKII, *Topological Methods in the Theory of Nonlinear Integral Equations*, Macmillan, New York, 1964.

[10] ———, *Positive Solutions of Operator Equations*, P. Noordhoff, Groningen, the Netherlands, 1964.

[11] R. D. NUSSBAUM, *A global bifurcation theorem with applications to functional differential equations*, J. Functional Anal., 19 (1975), pp. 319–338.

[12] ———, *Asymptotic fixed point theorems for local condensing maps*, Math. Ann., 191 (1971), pp. 181–195.

[13] ———, *The fixed point index for local condensing maps*, Ann. Math. Pura Appl., 89 (1971), pp. 217–258.

[14] ———, *Periodic solutions of some nonlinear, autonomous functional differential equations. II*, J. Differential Equations, 14 (1973), pp. 360–394.

[15] ———, *Periodic solutions of some nonlinear integral equations*, Dynamical Systems, A. Bednarek and L. Cesari, eds., Academic Press, New York, 1977, pp. 221–251.

[16] P. H. RABINOWITZ, *Some global results for nonlinear eigenvalue problems*, J. Functional Analysis, 7 (1971), pp. 487–513.

[17] H. SCHAEFER, *Topological Vector Spaces*, Springer-Verlag, New York, 1971.

[18] H. SMITH, *Periodic solutions for a class of epidemic equations*, preprint.

[19] ———, *On periodic solutions of delay integral equations modelling epidemics and population growth*, Ph.D. dissertation, Univ. of Iowa, Iowa City, May 1976.

[20] R. E. L. TURNER, *Transversality and cone maps*, Arch. Rational Mech. Anal., 58 (1975), pp. 151–179.

[21] P. WALTMAN, *Deterministic Threshold Models in the Theory of Epidemics*, Lecture Notes in Biomathematics, Springer-Verlag, New York, 1974.

SIAM J. MATH. ANAL.
Vol. 9, No. 2, April 1978
Copyright © 1978 Society for Industrial and Applied Mathematics

ON CERTAIN MULTIPLE INTEGRALS OCCURRING IN A WAVEGUIDE SCATTERING PROBLEM*

J. BOERSMA†

Abstract. Closed-form results are presented for some n-fold integrals where the integrand contains the exponential of a specific quadratic form in n variables. These integrals arise in the ray-optical analysis of reflection and diffraction problems for an open-ended parallel-plane waveguide. The results are obtained by three methods: the first method is elementary, the second method uses an integral equation which is solved by the Wiener–Hopf technique, and the third method is based on a probabilistic interpretation of the integrals.

1. Introduction. This paper deals with the evaluation of the n-fold integrals

$$I_{n,q}(\alpha)=\pi^{-n/2}\int_0^\infty\cdots\int_0^\infty x_1^q\exp\left[-\alpha x_1^2-2\sum_{m=2}^n x_m^2+2\sum_{m=1}^{n-1}x_mx_{m+1}\right]dx_1\cdots dx_n,\tag{1.1}$$
$$n=1,2,3,\cdots,\qquad q=0,1,2,\cdots,$$

$$J_{n,q}(\alpha)=\pi^{-n/2}\int_0^\infty\cdots\int_0^\infty x_1^q\exp\left[-\alpha x_1^2-2\sum_{m=2}^{n-1}x_m^2+2\sum_{m=1}^{n-2}x_mx_{m+1}-2x_{n-1}x_n-x_n^2\right]dx_1\cdots dx_n,\tag{1.2}$$
$$n=2,3,4,\cdots,\qquad q=0,1,2,\cdots,$$

where the integration extends over the orthant $x_m\geqq 0$, $m=1,2,\cdots,n$. By repeated application of the estimate

$$\tfrac{1}{2}\pi^{1/2}p^{-1/2}\exp(s^2/p)\leqq\int_0^\infty\exp[-px^2+2sx]\,dx\leqq\pi^{1/2}p^{-1/2}\exp(s^2/p),\tag{1.3}$$

valid for $p>0$, $s\geqq 0$, it is found that the integrals (1.1) and (1.2) converge if $\alpha>(n-1)/n$ and $\alpha>(n-2)/(n-1)$, respectively.

These integrals were encountered in the ray-optical analysis of (i) the reflection problem for a TM or a TE mode traveling toward the open end of a semi-infinite parallel-plane waveguide [6], [7], (ii) the diffraction problem for a plane wave normally incident on two nonstaggered parallel half-planes [14]. In the course of that analysis explicit results were needed for $I_{n,q}(\alpha)$, $J_{n,q}(\alpha)$ with $q=0$, 1 and $\alpha=1$ or $\alpha=2$. It is the purpose of this paper to provide such results, namely

$$I_{n,0}(2)=\frac{1}{(n+1)^{3/2}},\qquad I_{n,1}(2)=\frac{1}{4\pi^{1/2}}\sum_{m=1}^n\frac{1}{m^{3/2}(n+1-m)^{3/2}},\tag{1.4}$$

$$I_{n,0}(1)=\frac{(1/2)_n}{n!},\qquad I_{n,1}(1)=\frac{1}{2\pi^{1/2}}\sum_{m=0}^{n-1}\frac{(1/2)_m}{m!(n-m)^{1/2}},\tag{1.5}$$

$$J_{n,0}(2)=-\frac{(-1/2)_n}{n!},\qquad J_{n,1}(2)=-\frac{1}{2\pi^{1/2}}\sum_{m=0}^{n-1}\frac{(-1/2)_m}{m!(n-m)^{3/2}},\tag{1.6}$$

* Received by the editors June 4, 1976, and in revised form October 14, 1976.

† Department of Mathematics, Technological University, Eindhoven, the Netherlands.

$$(1.7)\quad J_{n,0}(1)=\frac{1}{2\pi(n-1)^{1/2}},\qquad J_{n,1}(1)=\frac{1}{8\pi^{1/2}}+\frac{1}{8\pi^{3/2}}\sum_{m=1}^{n-2}\frac{1}{m^{1/2}(n-1-m)^{1/2}},$$

where $(a)_n$ denotes Pochhammer's symbol defined by

$$(1.8)\qquad (a)_0=1,\quad (a)_n=a(a+1)\cdots(a+n-1),\qquad n=1,2,3,\cdots.$$

The integrals $I_{n,0}(2)$ and $I_{n,1}(2)$ were already evaluated [6, Appendix D] thus leading to (1.4). Only recently the author became aware of a previous evaluation of $I_{n,0}(2)$ by Anis and Lloyd [2] using essentially the same method as in [6, Appendix D]. In § 2 of this paper the remaining results (1.5)–(1.7) are derived by elementary methods that involve integration by parts and generating function techniques. For $q \geqq 2$ recurrence relations are presented for $I_{n,q}$, $J_{n,q}$, expressed in terms of the same functions with second subscripts $q-1$ and $q-2$.

In §§ 3.1, 3.2 the results (1.4)–(1.7) are rederived by a second and different approach. It is shown that the evaluation of $I_{n,q}$ and $J_{n,q}$ can be reduced to the solution of the integral equation

$$(1.9)\qquad \varphi(t)=f(t)+\frac{\lambda}{\pi^{1/2}}\int_0^\infty \exp[-(t-s)^2]\varphi(s)\,ds,$$

where $f(t)=e^{-t^2}$ and $|\lambda|<1$. The latter equation is solved by Fourier transformation and Wiener–Hopf technique. In § 3.3 we consider the related n-fold integral

$$(1.10)\qquad \mathcal{T}_n(t)=\pi^{-n/2}\int_0^\infty\cdots\int_0^\infty \exp\Big[2\,e^{-\pi i/4}tx_1-x_1^2-2\sum_{m=2}^{n}x_m^2 +2\sum_{m=1}^{n-1}x_mx_{m+1}\Big]\,dx_1\cdots dx_n,\qquad n=1,2,3,\cdots,$$

while $\mathcal{T}_0(t)=1$ by definition. As a side result of the previous analysis it is found that

$$(1.11)\qquad \sum_{n=0}^{\infty}\lambda^n\mathcal{T}_n(t)=\begin{cases}\exp\Big[\dfrac{1}{2\pi i}\displaystyle\int_{-\infty}^{\infty}\dfrac{\log(1-\lambda\,e^{-x^2})}{x-e^{\pi i/4}t}\,dx\Big], & t<0,\\[2ex] (1-\lambda\,e^{-it^2})^{-1}\exp\Big[\dfrac{1}{2\pi i}\displaystyle\int_{-\infty}^{\infty}\dfrac{\log(1-\lambda\,e^{-x^2})}{x-e^{\pi i/4}t}\,dx\Big], & t>0,\end{cases}$$

where $|\lambda|<1$. Then the right-hand side of (1.11) is expanded in a power series in powers of λ and it turns out that $\mathcal{T}_n(t)$ can be expressed in terms of Fresnel integrals. The result (1.11) is to be used in the ray-optical solution of the radiation problem for an incident mode traveling toward the open end of a semi-infinite parallel-plane waveguide [8].

The Wiener–Hopf solution of the integral equation (1.9) has been treated in the literature to some extent. Stewartson [18] solved both (1.9) and the associated homogeneous equation with $f(t)=0$, in the case when $\lambda=1$. As Stewartson points out, these integral equations arise in the evolution theory of comets and in some problems from fluid mechanics. Ghizzetti and Ossicini [11] studied the eigensolutions of the homogeneous equation when $\lambda>0$. A related integral equation with a shifted kernel $\exp[-\frac{1}{2}(t+\Delta-s)^2]$ was recently discussed by Atkinson [5] in connection with some inference and queuing problems.

In § 4 the integrals $I_{n,0}(\alpha)$ and $J_{n,0}(\alpha)$ with $\alpha=1$ or $\alpha=2$ are evaluated by a third, probabilistic method. It is shown that $I_{n,0}$ and $J_{n,0}$ can be interpreted in terms of the probability distribution of a sum of random variables which are independent and have the same normal distribution function. Then the explicit values (1.4)–(1.7) are recovered by means of theorems due to Sparre Andersen [1] and Spitzer [17]. In fact, the same probabilistic approach underlies a previous evaluation of $I_{n,0}(1)$ due to Anis and Lloyd [2], [3].

Integrals similar to (1.1), (1.2), but containing the exponential of a general quadratic form, occur in probability theory and statistics. For example, the so-called orthant probability for the multivariate normal distribution with zero means and variance-covariance matrix $\mathbf{V}$ is given by

$$\Phi_n(\mathbf{V}) = (2\pi)^{-n/2}|\mathbf{V}|^{-1/2}\int_0^\infty \cdots \int_0^\infty \exp[-\tfrac{1}{2}\mathbf{x}'\mathbf{V}^{-1}\mathbf{x}]\,dx_1\cdots dx_n, \tag{1.12}$$

where $\mathbf{x}' = (x_1, x_2, \cdots, x_n)$; see Ruben [16], Johnson and Kotz [12, Chap. 35]. According to [16], $\Phi_n(\mathbf{V})$ can be expressed in terms of the area of a certain simplex on the unit sphere in n-dimensional space. Such an expression is obtained by a suitable linear transformation of (1.12) which reduces $\mathbf{x}'\mathbf{V}^{-1}\mathbf{x}$ to a sum of squares. Then the domain of integration is transformed into a polyhedral cone in n-dimensional space, bounded by n hyperplanes through the origin, and the said simplex is the intersection of the cone and the unit sphere. Closed-form results for $\Phi_n(\mathbf{V})$ are readily obtained now in the cases $n = 1, 2, 3$. For $n > 3$, $\Phi_n(\mathbf{V})$ can no longer be expressed in terms of elementary functions; cf. [16, p. 171]. Various other methods for the evaluation of multinormal probabilities are reviewed in [12, Chap. 35]. It is remarked that none of the methods of this paper is applicable to the general integral (1.12). As for the integrals $I_{n,q}(\alpha)$, $J_{n,q}(\alpha)$, closed-form results valid for any α may be derived when $n \leqq 3$ by the geometrical approach as described above.

2. Evaluation by elementary means.

2.1. $I_{n,q}(2)$. The integrals $I_{n,0}(2)$ and $I_{n,1}(2)$ were already evaluated, see [6, Appendix D], [2]. Consider now $I_{n,q}(2)$ with $q \geqq 2$, as defined by (1.1), and replace the factor x_1^q in the integrand by

$$x_1^q = \frac{x_1^{q-1}}{2(n+1)}\left[n(4x_1 - 2x_2) + \sum_{m=2}^{n}(n+1-m)(-2x_{m-1} + 4x_m - 2x_{m+1})\right], \tag{2.1}$$

where $x_{n+1} = 0$ by definition. Then $I_{n,q}(2)$ can be expressed as a sum of integrals which permit integration by parts with respect to x_1 and explicit integration with respect to x_m, $m = 2, 3, \cdots, n$, respectively. The result comprises an n-fold integral which is recognized as $I_{n,q-2}(2)$, and a sum of $(n-1)$-fold integrals which can be expressed as products $I_{m-1,q-1}(2)I_{n-m,0}(2)$, $m = 2, 3, \cdots, n$. On substitution of the actual value of $I_{n-m,0}(2)$, we obtain the recurrence relation

$$I_{n,q}(2) = \frac{n(q-1)}{2(n+1)}I_{n,q-2}(2) + \frac{\pi^{-1/2}}{2(n+1)}\sum_{m=1}^{n-1}\frac{I_{m,q-1}(2)}{(n-m)^{1/2}}, \tag{2.2}$$

valid for $q \geqq 2$. The same method may be used for the reduction of the integral $I_{n,1}(2)$, yielding

$$I_{n,1}(2) = \frac{\pi^{-1/2}}{2(n+1)}\sum_{m=1}^{n}\frac{I_{m-1,0}(2)}{(n+1-m)^{1/2}} = \frac{1}{4\pi^{1/2}}\sum_{m=1}^{n}\frac{1}{m^{3/2}(n+1-m)^{3/2}}, \tag{2.3}$$

in accordance with (1.4). The relations (2.2) and (2.3) can be combined to the single recurrence relation

$$I_{n,q}(2)=\frac{n(q-1)}{2(n+1)}I_{n,q-2}(2)+\frac{\pi^{-1/2}}{2(n+1)}\sum_{m=0}^{n-1}\frac{I_{m,q-1}(2)}{(n-m)^{1/2}}, \tag{2.4}$$

valid for $q \geqq 1$, where it is understood that $(q-1)I_{n,q-2}(2)=0$ for $q=1$, and $I_{0,q}(2)=\delta_{q0}$ with $\delta_{00}=1$, $\delta_{q0}=0$ for $q\neq 0$ (Kronecker's symbol).

2.2. $I_{n,q}(1)$, $J_{n,q}(2)$. Consider first the integrals $I_{n,0}(1)$ and $J_{n,0}(2)$, as defined by (1.1) and (1.2). In the integral $J_{n,0}(2)$ we set

$$\int_0^\infty \exp[-2x_{n-1}x_n-x_n^2]\,dx_n=\pi^{1/2}\exp(x_{n-1}^2)-\int_0^\infty \exp[2x_{n-1}x_n-x_n^2]\,dx_n;$$

then it is obvious that

$$J_{n,0}(2)=I_{n-1,0}(1)-I_{n,0}(1), \qquad n\geqq 2. \tag{2.5}$$

A second relation between $I_{n,0}(1)$ and $J_{n,0}(2)$ is obtained by starting from the identity

$$\begin{aligned}\pi^{-n/2}\int_0^\infty\cdots\int_0^\infty\Big\{&(2x_1-2x_2)+\sum_{m=2}^{n-1}(-2x_{m-1}+4x_m-2x_{m+1})\\ &+(-2x_{n-1}+4x_n+2x_{n+1})-(2x_n+2x_{n+1})\Big\}\\ &\cdot\exp\Big[-x_1^2-2\sum_{m=2}^{n}x_m^2+2\sum_{m=1}^{n-1}x_mx_{m+1}-2x_nx_{n+1}\\ &\qquad -x_{n+1}^2\Big]\,dx_1\cdots dx_{n+1}=0,\end{aligned} \tag{2.6}$$

where $n\geqq 2$. Notice that the successive linear factors are just the derivatives of the exponent with respect to x_m, $m=1,2,\cdots,n+1$. Hence, the $(n+1)$-fold integral (2.6) can be rewritten as a sum of integrals which permit explicit integration with respect to x_m. Each of the resulting n-fold integrals is the product of an $(m-1)$-fold integral equal to $I_{m-1,0}(1)$, and an $(n+1-m)$-fold integral equal to $J_{n+1-m,0}(2)$. Thus we find

$$\sum_{m=0}^{n} I_{m,0}(1)J_{n-m,0}(2)=0, \qquad n\geqq 2, \tag{2.7}$$

where $I_{0,0}(1)=1$, $J_{0,0}(2)=-1$, $J_{1,0}(2)=\frac{1}{2}$ by definition. By a direct calculation from (1.1) it is found that $I_{1,0}(1)=\frac{1}{2}$, hence, (2.7) and (2.5) also hold for $n=1$.

In order to determine $I_{n,0}(1)$ and $J_{n,0}(2)$, we introduce the generating functions

$$A_0(\lambda)=\sum_{n=0}^{\infty} I_{n,0}(1)\lambda^n, \qquad B_0(\lambda)=\sum_{n=0}^{\infty} J_{n,0}(2)\lambda^n. \tag{2.8}$$

From (2.5) and (2.7) we then infer

$$B_0(\lambda)=-(1-\lambda)A_0(\lambda), \qquad A_0(\lambda)B_0(\lambda)=-1, \tag{2.9}$$

and consequently

$$A_0(\lambda)=(1-\lambda)^{-1/2}, \qquad B_0(\lambda)=-(1-\lambda)^{1/2}. \tag{2.10}$$

By expansion of $(1-\lambda)^{\pm 1/2}$ in binomial series we readily find $I_{n,0}(1)$ and $J_{n,0}(2)$, as stated in (1.5) and (1.6).

Consider next the integral $I_{n,q}(1)$ with $q \geqq 1$, as defined by (1.1), and replace the factor x_1^q in the integrand by

$$x_1^q = \tfrac{1}{2} x_1^{q-1}\left[n(2x_1 - 2x_2) + \sum_{m=2}^{n} (n+1-m)(-2x_{m-1} + 4x_m - 2x_{m+1}) \right], \tag{2.11}$$

where $x_{n+1} = 0$ by definition. Proceeding as in § 2.1, we are led to the recurrence relation

$$I_{n,q}(1) = \tfrac{1}{2} n(q-1) I_{n,q-2}(1) + \frac{1}{2\pi^{1/2}} \sum_{m=0}^{n-1} \frac{I_{m,q-1}(1)}{(n-m)^{1/2}}, \tag{2.12}$$

valid for $q \geqq 1$, where $(q-1)I_{n,q-2}(1) = 0$ for $q = 1$, and $I_{0,q}(1) = \delta_{q0}$ by definition. The present relation was also established by Anis [4] in the same manner. A similar recurrence relation for $J_{n,q}(2)$ is obtained by setting

$$\begin{aligned} x_1^q = \tfrac{1}{2} x_1^{q-1}\Big[& (4x_1 - 2x_2) + \sum_{m=2}^{n-2} (-2x_{m-1} + 4x_m - 2x_{m+1}) \\ & + (-2x_{n-2} + 4x_{n-1} + 2x_n) - (2x_{n-1} + 2x_n) \Big] \end{aligned} \tag{2.13}$$

in the defining integral (1.2). Thus we find

$$J_{n,q}(2) = \tfrac{1}{2}(q-1) J_{n,q-2}(2) - \frac{1}{2\pi^{1/2}} \sum_{m=0}^{n-1} \frac{(-1/2)_m}{m!} I_{n-1-m,q-1}(2), \tag{2.14}$$

valid for $q \geqq 1$, where $(q-1)J_{n,q-2}(2) = 0$ for $q = 1$, and $I_{0,q}(2) = \delta_{q0}$ by definition. For $q = 1$ the recurrence relations (2.12) and (2.14) provide the explicit values of $I_{n,1}(1)$ and $J_{n,1}(2)$, as stated in (1.5) and (1.6).

For later use we introduce the generating function

$$A_q(\lambda) = \sum_{n=0}^{\infty} I_{n,q}(1) \lambda^n, \qquad q = 0, 1, 2, \cdots. \tag{2.15}$$

Then (2.12) can be reduced to a recurrence relation for $A_q(\lambda)$, viz.,

$$A_q(\lambda) = \frac{1}{2}(q-1)\lambda A'_{q-2}(\lambda) + \frac{1}{2\pi^{1/2}} A_{q-1}(\lambda) L(\lambda), \qquad q \geqq 1, \tag{2.16}$$

where a prime denotes differentiation with respect to λ and

$$L(\lambda) = \sum_{n=1}^{\infty} \frac{\lambda^n}{n^{1/2}}. \tag{2.17}$$

Starting from $A_0(\lambda) = (1-\lambda)^{-1/2}$, we have

$$A_1(\lambda) = \frac{1}{2\pi^{1/2}} (1-\lambda)^{-1/2} L(\lambda), \tag{2.18}$$

$$A_2(\lambda) = \frac{1}{4} \lambda (1-\lambda)^{-3/2} + \frac{1}{4\pi} (1-\lambda)^{-1/2} L^2(\lambda), \tag{2.19}$$

and so on; in principle the functions $A_q(\lambda)$ are completely determined.

2.3. $J_{n,q}(1)$. Starting from (1.2) with $\alpha=1$, we perform an integration by parts with respect to x_1, yielding

$$J_{n,q}(1)=\frac{\pi^{-n/2}}{q+1}\int_0^\infty\cdots\int_0^\infty x_1^{q+1}(2x_1-2x_2)\exp\Big[-x_1^2-2\sum_{m=2}^{n-1}x_m^2 +2\sum_{m=1}^{n-2}x_mx_{m+1}-2x_{n-1}x_n-x_n^2\Big]\,dx_1\cdots dx_n, \tag{2.20}$$

where it is supposed that $n\geqq 3$. In the latter integral the factor $2x_1-2x_2$ is replaced by

$$2x_1-2x_2=-\sum_{m=2}^{n-2}(-2x_{m-1}+4x_m-2x_{m+1}) -(-2x_{n-2}+4x_{n-1}+2x_n)+(2x_{n-1}+2x_n). \tag{2.21}$$

Then $J_{n,q}(1)$ becomes a sum of integrals which permit explicit integration with respect to x_m, $m=2,3,\cdots,n$. Proceeding as before, we find

$$J_{n,q}(1)=\frac{\pi^{-1/2}}{q+1}\sum_{m=0}^{n-2}\frac{(-1/2)_m}{m!}I_{n-1-m,q+1}(1), \tag{2.22}$$

valid for $n\geqq 3$. In a similar manner it can be verified that (2.22) holds true also for $n=2$. Thus $J_{n,q}(1)$ has been expressed in terms of the integrals $I_{m,q+1}(1)$ which are known from § 2.2.

In order to explicitly evaluate $J_{n,0}(1)$ and $J_{n,1}(1)$, we introduce the generating function

$$C_q(\lambda)=\sum_{n=2}^\infty J_{n,q}(1)\lambda^n. \tag{2.23}$$

Then it follows from (2.22) that

$$C_q(\lambda)=\frac{\pi^{-1/2}}{q+1}\lambda(1-\lambda)^{1/2}A_{q+1}(\lambda) \tag{2.24}$$

where $A_{q+1}(\lambda)$ is defined by (2.15). Referring to (2.18), (2.19), we thus find

$$C_0(\lambda)=(1/2\pi)\lambda L(\lambda), \tag{2.25}$$

$$C_1(\lambda)=\frac{1}{8\pi^{1/2}}\lambda^2(1-\lambda)^{-1}+\frac{1}{8\pi^{3/2}}\lambda L^2(\lambda). \tag{2.26}$$

By expansion of these functions the results (1.7) for $J_{n,0}(1)$ and $J_{n,1}(1)$ are readily established.

3. Evaluation by integral equations.

3.1. $I_{n,q}(2), I_{n,q}(1)$. Let the functions $\varphi_n(t)$, t real, $n=0,1,2,\cdots$, be defined by

$$\varphi_0(t)=e^{-t^2},$$

$$\varphi_n(t)=\pi^{-n/2}e^{-t^2}\int_0^\infty\cdots\int_0^\infty\exp\Big[2tx_1-2\sum_{m=1}^n x_m^2+2\sum_{m=1}^{n-1}x_mx_{m+1}\Big]\,dx_1\cdots dx_n, \quad n=1,2,3,\cdots. \tag{3.1}$$

Then it is easily seen from (1.1) that

$$I_{n,q}(2)=2^{-q}\Big(\frac{d}{dt}\Big)^q[e^{t^2}\varphi_n(t)]\Big|_{t=0},\qquad I_{n,q}(1)=\pi^{-1/2}\int_0^\infty t^q\varphi_{n-1}(t)\,dt. \tag{3.2}$$

By repeated application of (1.3), one is led to the estimate

$$
\begin{aligned}
0 \leqq \varphi_n(t) &\leqq \pi^{-1/2} n^{-1/2} e^{-t^2} \int_0^\infty \exp\left[2tx_1 - \frac{n+1}{n} x_1^2\right] dx_1 \\
&\leqq \begin{cases} (n+1)^{-1/2} \exp[-t^2/(n+1)], & t \geqq 0, \\ \tfrac{1}{2}(n+1)^{-1/2} \exp(-t^2), & t \leqq 0. \end{cases}
\end{aligned} \tag{3.3}
$$

The functions $\varphi_n(t)$ are connected through the recurrence relation

$$
\varphi_n(t) = \pi^{-1/2} \int_0^\infty \exp[-(t-s)^2] \varphi_{n-1}(s)\, ds, \qquad n \geqq 1. \tag{3.4}
$$

We now introduce the generating function

$$
\varphi(t) = \sum_{n=0}^{\infty} \lambda^n \varphi_n(t), \qquad |\lambda| < 1; \tag{3.5}
$$

then, in view of (3.3), (3.4), the latter series converges and is precisely the Neumann series associated with the integral equation

$$
\varphi(t) = e^{-t^2} + \frac{\lambda}{\pi^{1/2}} \int_0^\infty \exp[-(t-s)^2] \varphi(s)\, ds. \tag{3.6}
$$

Furthermore, it follows from (3.3) that

$$
\varphi(t) = O(1), \qquad t \geqq 0; \quad \varphi(t) = O(e^{-t^2}), \qquad t \leqq 0. \tag{3.7}
$$

(By a more careful analysis the first result can even be improved to $\varphi(t) = O(e^{-\beta t})$ as $t \to \infty$, where β will be specified below; however, we do not need this sharper estimate.)

The integral equation (3.6) is solved by Fourier transformation and Wiener–Hopf technique (cf. Noble [15]). We introduce the Fourier transforms

$$
\Phi_+(w) = \int_0^\infty \varphi(t)\, e^{iwt}\, dt, \qquad \Phi_-(w) = \int_{-\infty}^0 \varphi(t)\, e^{iwt}\, dt, \tag{3.8}
$$

where w is a complex variable. Then the estimates (3.7) imply that $\Phi_+(w)$ is regular in the upper half-plane $\operatorname{Im} w > 0$, while $\Phi_-(w)$ is an integral function. Under Fourier transformation the integral equation (3.6) reduces to

$$
\Phi_+(w) + \Phi_-(w) = \pi^{1/2} e^{-w^2/4} + \lambda\, e^{-w^2/4} \Phi_+(w), \qquad \operatorname{Im} w > 0,
$$

or equivalently

$$
(1 - \lambda\, e^{-w^2/4}) \left[\Phi_+(w) + \frac{\pi^{1/2}}{\lambda}\right] + \left[\Phi_-(w) - \frac{\pi^{1/2}}{\lambda}\right] = 0, \qquad \operatorname{Im} w > 0. \tag{3.9}
$$

Before going on we observe that the factor $1 - \lambda\, e^{-w^2/4}$ vanishes when $w = 2(\log \lambda)^{1/2}$. The zeros closest to the real axis have imaginary parts $\pm\beta$ where $\beta = 2|\operatorname{Im}(\log \lambda)^{1/2}|$ with the principal value of $\log \lambda$ to be taken. Thus, $1 - \lambda\, e^{-w^2/4} \neq 0$ in the strip $-\beta < \operatorname{Im} w < \beta$. Then by means of (3.9), extended to $\operatorname{Im} w > -\beta$, $\Phi_+(w)$ may be analytically continued into the upper half-plane $\operatorname{Im} w > -\beta$.

The functional equation (3.9) is now solved by the standard Wiener–Hopf procedure. The key step in this procedure is the factorization of $1 - \lambda\, e^{-w^2/4}$ into

$$
1 - \lambda\, e^{-w^2/4} = K_+(w)/K_-(w), \qquad -\beta < \operatorname{Im} w < \beta, \tag{3.10}
$$

such that $K_+(w)$ is regular and nonzero in $\operatorname{Im} w > -\beta$, and $K_-(w)$ is regular and nonzero in $\operatorname{Im} w < \beta$. This factorization can be accomplished by means of Noble [15, § 1.3, Thm. C], yielding

$$K_+(w) = \exp\left[\frac{1}{2\pi i}\int_{-\infty+ia}^{\infty+ia} \frac{\log(1-\lambda e^{-z^2/4})}{z-w}\,dz\right], \tag{3.11}$$

$$K_-(w) = \exp\left[\frac{1}{2\pi i}\int_{-\infty+ib}^{\infty+ib} \frac{\log(1-\lambda e^{-z^2/4})}{z-w}\,dz\right], \tag{3.12}$$

where a, b are any numbers subject to $-\beta < a < b < \beta$, and the logarithm stands for its principal value. Using (3.10), we rearrange (3.9) as

$$K_+(w)\left[\Phi_+(w) + \frac{\pi^{1/2}}{\lambda}\right] = -K_-(w)\left[\Phi_-(w) - \frac{\pi^{1/2}}{\lambda}\right], \qquad -\beta < \operatorname{Im} w < \beta. \tag{3.13}$$

Then the functions on the left-hand side of (3.13) are regular in the upper half-plane $\operatorname{Im} w > -\beta$, and the functions on the right-hand side are regular in the lower half-plane $\operatorname{Im} w < \beta$. Hence, by analytic continuation both sides of (3.13) must equal an integral function $P(w)$, say. From (3.11), (3.12) it is obvious that $K_\pm(w) \to 1$ as $|w| \to \infty$, $\operatorname{Im} w \gtrless \mp\beta$; likewise, $\Phi_\pm(w) \to 0$ as $|w| \to \infty$, $\operatorname{Im} w \gtrless \mp\beta$, according to the Riemann–Lebesgue lemma. Thus $P(w) \to \pi^{1/2}/\lambda$ as $|w| \to \infty$, and consequently $P(w) = \pi^{1/2}/\lambda$ by Liouville's theorem. Then the solution for $\Phi_+(w)$ is easily obtained from the left-hand side of (3.13), viz.,

$$\begin{aligned} \Phi_+(w) &= \frac{\pi^{1/2}}{\lambda}\left\{\frac{1}{K_+(w)} - 1\right\} \\ &= \frac{\pi^{1/2}}{\lambda}\left\{\exp\left[-\frac{1}{2\pi i}\int_{-\infty}^{\infty} \frac{\log(1-\lambda e^{-z^2/4})}{z-w}\,dz\right] - 1\right\}, \qquad \operatorname{Im} w > 0, \end{aligned} \tag{3.14}$$

where the path of integration has been chosen along the real axis. Finally, the original function $\varphi(t)$ is found by inverse Fourier transformation of $\Phi_+(w)$, viz.,

$$\varphi(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi_+(w)\, e^{-itw}\,dw, \qquad t > 0. \tag{3.15}$$

The solution for $\varphi(t)$ thus determined is of a rather complicated form. However, it follows from (3.2), (3.5), (3.8) that

$$\sum_{n=0}^{\infty} I_{n,q}(2)\lambda^n = 2^{-q}\left(\frac{d}{dt}\right)^q [e^{t^2}\varphi(t)]\Big|_{t=0}, \tag{3.16}$$

$$\sum_{n=1}^{\infty} I_{n,q}(1)\lambda^n = \frac{\lambda}{\pi^{1/2}}\int_0^{\infty} t^q \varphi(t)\,dt = \frac{\lambda i^{-q}}{\pi^{1/2}}\Phi_+^{(q)}(0), \tag{3.17}$$

so the required integrals $I_{n,q}(2)$ and $I_{n,q}(1)$ are completely determined by the derivatives $\varphi^{(m)}(0)$, $m = 0, 1, \cdots, q$, and $\Phi_+^{(q)}(0)$ only. The derivatives $\varphi^{(m)}(0)$ are readily obtained from the asymptotic expansion of $\Phi_+(w)$ as $|w| \to \infty$, $\operatorname{Im} w > 0$. In fact, starting from the Taylor series

$$\varphi(t) = \sum_{m=0}^{\infty} \frac{\varphi^{(m)}(0)}{m!} t^m, \tag{3.18}$$

one has by Watson's lemma (see e.g. Erdélyi [9, § 2.2])

$$(3.19)\qquad \Phi_+(w) \sim \sum_{m=0}^{\infty} i^{m+1}\varphi^{(m)}(0)w^{-m-1}, \qquad |w|\to\infty, \quad \delta \leqq \arg w \leqq \pi-\delta,$$

for any positive δ. Thus the integrals $I_{n,q}(2)$ and $I_{n,q}(1)$ can be determined from $\Phi_+(w)$ only. We shall now evaluate the integrals in the two cases $q=0$ and $q=1$.

The asymptotic expansion of $\Phi_+(w)$ is easily obtained from (3.14), viz.,

$$(3.20)\qquad \begin{aligned}\Phi_+(w) = \frac{\pi^{1/2}}{\lambda}\Big\{&\frac{1}{2\pi i}\int_{-\infty}^{\infty} \log(1-\lambda\, e^{-z^2/4})\, dz\, w^{-1} \\ &-\frac{1}{8\pi^2}\Big[\int_{-\infty}^{\infty} \log(1-\lambda\, e^{-z^2/4})\, dz\Big]^2 w^{-2} + O(w^{-3})\Big\}\end{aligned}$$

as $|w|\to\infty$, Im $w>0$. Compare the latter expansion to (3.19); then it is found that

$$(3.21)\qquad \begin{aligned}\varphi(0) &= -\frac{\pi^{1/2}}{2\pi\lambda}\int_{-\infty}^{\infty} \log(1-\lambda\, e^{-z^2/4})\, dz \\ &= \frac{\pi^{1/2}}{2\pi\lambda}\sum_{n=1}^{\infty}\frac{\lambda^n}{n}\int_{-\infty}^{\infty} e^{-nz^2/4}\, dz = \sum_{n=0}^{\infty}\frac{\lambda^n}{(n+1)^{3/2}},\end{aligned}$$

$$(3.22)\qquad \tfrac{1}{2}\varphi'(0) = \frac{\lambda}{4\pi^{1/2}}[\varphi(0)]^2 = \frac{\lambda}{4\pi^{1/2}}\Big[\sum_{n=0}^{\infty}\frac{\lambda^n}{(n+1)^{3/2}}\Big]^2.$$

In view of (3.16), the present results immediately yield $I_{n,0}(2)$ and $I_{n,1}(2)$, their values being given by (1.4).

Next we determine $\Phi_+(0)$ from (3.14) by taking the limit when $w\to 0$ from the upper side Im $w>0$. By Plemelj's formulae we have

$$(3.23)\qquad \frac{\lambda}{\pi^{1/2}}\Phi_+(0) = \exp[-\tfrac{1}{2}\log(1-\lambda)] - 1 = (1-\lambda)^{-1/2} - 1 = \sum_{n=1}^{\infty}\frac{(1/2)_n}{n!}\lambda^n,$$

which should be compared to (3.17). Then the result (1.5) for $I_{n,0}(1)$ is obvious. From (3.14) the derivative $\Phi'_+(w)$ is found to be

$$(3.24)\quad \Phi'_+(w) = \Big[\Phi_+(w) + \frac{\pi^{1/2}}{\lambda}\Big]\Big(-\frac{\lambda}{4\pi i}\Big)\int_{-\infty}^{\infty}\frac{z\, e^{-z^2/4}}{1-\lambda\, e^{-z^2/4}}\frac{dz}{z-w}, \qquad \text{Im } w>0.$$

Then again by Plemelj's formulae we have

$$(3.25)\qquad \begin{aligned}\frac{\lambda i^{-1}}{\pi^{1/2}}\Phi'_+(0) &= \Big[\frac{\lambda}{\pi^{1/2}}\Phi_+(0) + 1\Big]\frac{\lambda}{4\pi}\int_{-\infty}^{\infty}\frac{e^{-z^2/4}}{1-\lambda\, e^{-z^2/4}}\, dz \\ &= \frac{\lambda}{2\pi^{1/2}}\sum_{n=0}^{\infty}\frac{(1/2)_n}{n!}\lambda^n \sum_{n=0}^{\infty}\frac{\lambda^n}{(n+1)^{1/2}},\end{aligned}$$

from which the result (1.5) for $I_{n,1}(1)$ is easily recovered.

In principle, the integrals $I_{n,q}(2)$ and $I_{n,q}(1)$ with $q\geqq 2$ can be evaluated in the same manner. In addition, one may establish recurrence relations for the integrals. However, these recurrence relations turn out to be more complicated than the ones derived in §§ 2.1, 2.2. Therefore we shall not pursue this matter.

3.2. $J_{n,q}(2)$, $J_{n,q}(1)$. The approach is highly similar to that of § 3.1. Let the functions $\psi_n(t)$, t real, $n=1, 2, 3, \cdots$, be defined by

$$\psi_1(t)=\tfrac{1}{2}\operatorname{erfc} t=\pi^{-1/2}\int_t^\infty e^{-x^2}\,dx,$$

(3.26)
$$\psi_n(t)=\pi^{-n/2}e^{-t^2}\int_0^\infty\cdots\int_0^\infty \exp\Big[2tx_1-2\sum_{m=1}^{n-1}x_m^2+2\sum_{m=1}^{n-2}x_mx_{m+1} -2x_{n-1}x_n-x_n^2\Big]\,dx_1\cdots dx_n,\qquad n=2,3,4,\cdots.$$

Then it is easily seen from (1.2) that

$$J_{n,q}(2)=2^{-q}\Big(\frac{d}{dt}\Big)^q[e^{t^2}\psi_n(t)]\Big|_{t=0},\qquad J_{n,q}(1)=\pi^{-1/2}\int_0^\infty t^q\psi_{n-1}(t)\,dt. \tag{3.27}$$

The inner x_n integral in (3.26) can be estimated in an obvious manner, thus leading to

$$0\leqq\psi_n(t)\leqq\tfrac{1}{2}\varphi_{n-1}(t)\leqq\begin{cases}\tfrac{1}{2}n^{-1/2}e^{-t^2/n}, & t\geqq 0,\\ \tfrac{1}{4}n^{-1/2}e^{-t^2}, & t\leqq 0,\end{cases}\qquad n\geqq 2, \tag{3.28}$$

on account of (3.3). The functions $\psi_n(t)$ are connected through the recurrence relation

$$\psi_n(t)=\pi^{-1/2}\int_0^\infty \exp[-(t-s)^2]\psi_{n-1}(s)\,ds,\qquad n\geqq 2. \tag{3.29}$$

As we did in (3.5), we introduce the generating function

$$\psi(t)=\sum_{n=1}^\infty \lambda^n\psi_n(t),\qquad |\lambda|<1; \tag{3.30}$$

then, in view of (3.29), the latter series is the Neumann series associated with the integral equation

$$\psi(t)=\tfrac{1}{2}\lambda\operatorname{erfc} t+\frac{\lambda}{\pi^{1/2}}\int_0^\infty \exp[-(t-s)^2]\psi(s)\,ds. \tag{3.31}$$

Furthermore, by use of (3.28) it can be shown that $\psi(t)\to 0$ as $t\to\infty$.

The integral equation (3.31) can again be solved by Fourier transformation and Wiener–Hopf technique. However, a simpler way out is to differentiate (3.31) with respect to t followed by an integration by parts in the integral term, thus yielding

$$\psi'(t)=\frac{\lambda}{\pi^{1/2}}[\psi(0)-1]e^{-t^2}+\frac{\lambda}{\pi^{1/2}}\int_0^\infty \exp[-(t-s)^2]\psi'(s)\,ds. \tag{3.32}$$

The latter integral equation is of the same form as (3.6), hence, its solution is given by

$$\psi'(t)=\frac{\lambda}{\pi^{1/2}}[\psi(0)-1]\varphi(t). \tag{3.33}$$

By integration of (3.33) over $[0,\infty)$, we have

(3.34)
$$\begin{aligned}-\psi(0)&=\frac{\lambda}{\pi^{1/2}}[\psi(0)-1]\int_0^\infty \varphi(t)\,dt\\ &=\frac{\lambda}{\pi^{1/2}}[\psi(0)-1]\Phi_+(0)=[\psi(0)-1][(1-\lambda)^{-1/2}-1],\end{aligned}$$

where $\Phi_+(0)$ was taken from (3.23). Thus we find

$$\psi(0)=1-(1-\lambda)^{1/2}, \tag{3.35}$$

$$\psi'(t)=-\pi^{-1/2}\lambda(1-\lambda)^{1/2}\varphi(t), \tag{3.36}$$

and $\psi(t)$ is completely determined in this manner.

We now turn to the evaluation of the integrals $J_{n,q}(2)$ and $J_{n,q}(1)$. It follows from (3.27), (3.30) that

$$\sum_{n=1}^{\infty} J_{n,q}(2)\lambda^n = 2^{-q}\left(\frac{d}{dt}\right)^q [e^{t^2}\psi(t)]\Big|_{t=0}, \tag{3.37}$$

$$\sum_{n=2}^{\infty} J_{n,q}(1)\lambda^n = \frac{\lambda}{\pi^{1/2}}\int_0^\infty t^q\psi(t)\,dt. \tag{3.38}$$

In the cases $q=0$, $q=1$, the right-hand side of (3.37) reduces to

$$\psi(0)=1-(1-\lambda)^{1/2}=-\sum_{n=1}^{\infty}\frac{(-1/2)_n}{n!}\lambda^n, \tag{3.39}$$

$$\tfrac{1}{2}\psi'(0)=-\tfrac{1}{2}\pi^{-1/2}\lambda(1-\lambda)^{1/2}\varphi(0)=-\frac{\lambda}{2\pi^{1/2}}\sum_{n=0}^{\infty}\frac{(-1/2)_n}{n!}\lambda^n\sum_{n=0}^{\infty}\frac{\lambda^n}{(n+1)^{3/2}}, \tag{3.40}$$

where $\varphi(0)$ was quoted from (3.21). Then it is easily recognized that $J_{n,0}(2)$ and $J_{n,1}(2)$ are given by (1.6). In the same manner one may evaluate $J_{n,q}(2)$ when $q \geqq 2$. Consider next (3.38) where the right-hand side is reduced through an integration by parts. By replacing $\psi'(t)$ by (3.36), we obtain

$$\begin{aligned}\sum_{n=2}^{\infty} J_{n,q}(1)\lambda^n &= \frac{\pi^{-1}}{q+1}\lambda^2(1-\lambda)^{1/2}\int_0^\infty t^{q+1}\varphi(t)\,dt\\ &= \frac{\pi^{-1/2}}{q+1}\lambda\sum_{n=0}^{\infty}\frac{(-1/2)_n}{n!}\lambda^n\sum_{n=1}^{\infty} I_{n,q+1}(1)\lambda^n,\end{aligned} \tag{3.41}$$

on account of (3.17). By equating the coefficients of corresponding powers of λ in (3.41), we re-obtain the recurrence relation (2.22). As shown at the end of § 2.3, the latter relation readily yields the explicit values of $J_{n,0}(1)$ and $J_{n,1}(1)$.

3.3. $\mathscr{T}_n(t)$. We consider the n-fold integrals $\mathscr{T}_n(t)$, t real, $n=0,1,2,\cdots$, defined by

$$\begin{aligned}&\mathscr{T}_0(t)=1,\\ &\mathscr{T}_n(t)=\pi^{-n/2}\int_0^\infty\cdots\int_0^\infty \exp\left[2\,e^{-\pi i/4}tx_1-x_1^2-2\sum_{m=2}^{n}x_m^2\right.\\ &\qquad\qquad\left.+2\sum_{m=1}^{n-1}x_mx_{m+1}\right]dx_1\cdots dx_n,\\ &\qquad\qquad n=1,2,3,\cdots\end{aligned} \tag{3.42}$$

These integrals can be expressed in terms of the functions φ_n as given by (3.1), viz.,

$$\mathscr{T}_n(t)=\pi^{-1/2}\int_0^\infty \exp[2\,e^{-\pi i/4}tx]\varphi_{n-1}(x)\,dx, \qquad n\geqq 1. \tag{3.43}$$

Then by means of (3.3) one has the estimate

$$|\mathcal{T}_n(t)| \leqq (\pi n)^{-1/2} \int_0^\infty \exp\left[2^{1/2}tx - \frac{x^2}{n}\right] dx \leqq \begin{cases} e^{nt^2/2}, & t \geqq 0, \\ \frac{1}{2}, & t \leqq 0. \end{cases} \tag{3.44}$$

We now introduce the generating function

$$G(\lambda, t) = \sum_{n=0}^{\infty} \lambda^n \mathcal{T}_n(t), \tag{3.45}$$

where λ is a complex variable. In view of (3.44), the latter series certainly converges for $|\lambda| < \exp(-t^2/2)$ when $t \geqq 0$, and for $|\lambda| < 1$ when $t \leqq 0$. Replace $\mathcal{T}_n(t)$ by (3.43), then $G(\lambda, t)$ reduces to

$$G(\lambda, t) = 1 + \frac{\lambda}{\pi^{1/2}} \int_0^\infty \exp[2\, e^{-\pi i/4} tx] \varphi(x)\, dx = 1 + \frac{\lambda}{\pi^{1/2}} \Phi_+(-2\, e^{\pi i/4} t), \tag{3.46}$$

on account of (3.5), (3.8). The Fourier transform $\Phi_+(w)$ was determined in § 3.1—see (3.14); it was also found that $\Phi_+(w)$ is regular in the upper half-plane $\operatorname{Im} w > -\beta$, where $\beta = 2|\operatorname{Im}(\log \lambda)^{1/2}|$. For the present purpose we rewrite the solution (3.14) with the path of integration shifted to $\operatorname{Im} w = a$ where a is any number such that $-\beta < a \leqq 0$. Then we have, under an obvious change of variable,

$$1 + \frac{\lambda}{\pi^{1/2}} \Phi_+(-2\, e^{\pi i/4} t) = \exp\left[\frac{1}{2\pi i} \int_{-\infty - ia/2}^{\infty - ia/2} \frac{\log(1 - \lambda\, e^{-z^2})}{z - e^{\pi i/4} t}\, dz\right], \qquad t < -2^{-1/2} a,$$

or equivalently,

$$1 + \frac{\lambda}{\pi^{1/2}} \Phi_+(-2\, e^{\pi i/4} t) = \begin{cases} \exp\left[\dfrac{1}{2\pi i} \displaystyle\int_{-\infty}^{\infty} \frac{\log(1 - \lambda\, e^{-x^2})}{x - e^{\pi i/4} t}\, dx\right], & t < 0, \\ (1 - \lambda\, e^{-it^2})^{-1} \exp\left[\dfrac{1}{2\pi i} \displaystyle\int_{-\infty}^{\infty} \frac{\log(1 - \lambda\, e^{-x^2})}{x - e^{\pi i/4} t}\, dx\right], & 0 < t < 2^{-1/2}\beta. \end{cases} \tag{3.47}$$

Notice that the inequality $0 < t < 2^{-1/2}\beta$ is certainly satisfied when $t > 0$, $|\lambda| < \exp(-t^2/2)$. For fixed t the right-hand side of (3.47) is a regular function of λ in the region $|\lambda| < 1$. Hence, its Taylor series, that is the series (3.45), will be convergent when $|\lambda| < 1$. Thus we obtain the final result (1.11) for the generating function of the integrals $\mathcal{T}_n(t)$.

Starting from (1.11), we shall express $\mathcal{T}_n(t)$ in terms of Fresnel integrals F, generally defined by

$$F(t) = \pi^{-1/2}\, e^{-\pi i/4}\, e^{-it^2} \int_{-\infty}^{t} e^{is^2}\, ds. \tag{3.48}$$

To that purpose, the exponent in the right-hand side of (1.11) is expanded in a power series in powers of λ. Then the coefficient of λ^n, $n = 1, 2, 3, \cdots$, can be reduced to

$$\frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{e^{-nx^2}}{x - e^{\pi i/4} t}\, dx = \begin{cases} -F(n^{1/2} t), & t < 0, \\ e^{-int^2} - F(n^{1/2} t), & t > 0, \end{cases} \tag{3.49}$$

according to a well-known integral representation for the Fresnel integral. On substitution of the latter result, we find

$$(3.50)\qquad \sum_{n=0}^{\infty} \lambda^n \mathcal{T}_n(t) = \exp\left[\sum_{n=1}^{\infty} \frac{\lambda^n}{n} F(n^{1/2}t)\right],$$

valid for $|\lambda| < 1$ and all t. Then, by equating the coefficients of corresponding powers of λ in (3.50), we have

$$(3.51)\qquad \begin{aligned} &\mathcal{T}_0(t) = 1, \qquad \mathcal{T}_1(t) = F(t),\\ &\mathcal{T}_2(t) = \tfrac{1}{2}F(2^{1/2}t) + \tfrac{1}{2}F^2(t),\\ &\mathcal{T}_3(t) = \tfrac{1}{3}F(3^{1/2}t) + \tfrac{1}{2}F(2^{1/2}t)F(t) + \tfrac{1}{6}F^3(t), \end{aligned}$$

and so on. In addition to these explicit results, differentiation of (3.50) with respect to λ yields

$$(3.52)\qquad \sum_{n=1}^{\infty} n\lambda^{n-1}\mathcal{T}_n(t) = \sum_{n=0}^{\infty} \lambda^n \mathcal{T}_n(t) \sum_{n=1}^{\infty} \lambda^{n-1} F(n^{1/2}t),$$

from which we derive the simple recurrence relation

$$(3.53)\qquad \mathcal{T}_n(t) = \frac{1}{n}\sum_{m=0}^{n-1} \mathcal{T}_m(t) F((n-m)^{1/2}t), \qquad n \geqq 1.$$

It is clear that the integrals $\mathcal{T}_n(t)$ are completely determined by (3.53) and the initial value $\mathcal{T}_0(t) = 1$.

4. Evaluation by probabilistic means.

4.1. $I_{n,0}(1)$, $I_{n,0}(2)$. Let the functions $F_n(t)$, t real, $n = 1, 2, 3, \cdots$, be defined by

$$(4.1)\quad F_n(t) = \pi^{-n/2} \int_0^{\infty} \cdots \int_0^{\infty} \int_t^{\infty} \exp\left[-2\sum_{m=1}^{n-1} x_m^2 + 2\sum_{m=1}^{n-1} x_m x_{m+1} - x_n^2\right] dx_1 \cdots dx_n.$$

Here it is understood that the lower limit t pertains to the integration with respect to x_n, all other integrations having lower limits 0. It is easily seen from (1.1) that

$$(4.2)\qquad F_n(0) = I_{n,0}(1), \qquad -F_n'(0) = \pi^{-1/2} I_{n-1,0}(2).$$

Consider the exponent in (4.1) which is rewritten as

$$(4.3)\qquad 2\sum_{m=1}^{n-1} x_m^2 - 2\sum_{m=1}^{n-1} x_m x_{m+1} + x_n^2 = x_1^2 + \sum_{m=2}^{n} (x_m - x_{m-1})^2.$$

We introduce the new variables

$$(4.4)\qquad y_1 = x_1; \quad y_m = x_m - x_{m-1}, \qquad m = 2, 3, \cdots, n,$$

and conversely,

$$(4.5)\qquad x_m = \sum_{j=1}^{m} y_j, \qquad m = 1, 2, \cdots, n.$$

Then (4.1) transforms into

$$(4.6)\qquad F_n(t) = \pi^{-n/2} \int \underset{D_n}{\cdots} \int \exp\left[-\sum_{m=1}^{n} y_m^2\right] dy_1 \cdots dy_n,$$

where D_n is an n-dimensional domain given by

(4.7) $$D_n: \sum_{j=1}^{m} y_j \geqq 0, \quad m=1,2,\cdots,n-1; \qquad \sum_{j=1}^{n} y_j \geqq t.$$

The integral (4.6) admits of a simple probabilistic interpretation. Let $y_1, y_2, \cdots, y_n$ be independent random variables with a common normal density function $\pi^{-1/2} \exp(-y_m^2)$, $m=1,2,\cdots,n$, and let their partial sums be denoted by $S_m = \sum_{j=1}^{m} y_j$, $m=1,2,\cdots,n$. Then $F_n(t)$, as given by (4.6), is equal to the probability

$$P\{S_1 \geqq 0, \cdots, S_{n-1} \geqq 0, S_n \geqq t\}.$$

In particular we now have, from (4.2),

(4.8) $$I_{n,0}(1) = P\{S_1 \geqq 0, S_2 \geqq 0, \cdots, S_n \geqq 0\},$$

(4.9) $$I_{n-1,0}(2) = \pi^{1/2} p\{S_1 \geqq 0, \cdots, S_{n-1} \geqq 0, S_n = 0\},$$

where p denotes the probability density.

The probability (4.8) can be determined by means of the generating function relation

(4.10) $$1 + \sum_{n=1}^{\infty} P\{S_1 \geqq 0, S_2 \geqq 0, \cdots, S_n \geqq 0\}\lambda^n = \exp\left[\sum_{n=1}^{\infty} \frac{\lambda^n}{n} P\{S_n \geqq 0\}\right], \qquad |\lambda| < 1,$$

first proved by Sparre Andersen [1, Thm. 1]; for later, different proofs, see Spitzer [17, p. 330], Feller [10, § XII.7]. In the present case one has $P\{S_n \geqq 0\} = \frac{1}{2}$ for all n, thus leading to

(4.11) $$1 + \sum_{n=1}^{\infty} I_{n,0}(1)\lambda^n = \exp\left[\frac{1}{2} \sum_{n=1}^{\infty} \frac{\lambda^n}{n}\right] = \exp[-\tfrac{1}{2}\log(1-\lambda)] = (1-\lambda)^{-1/2},$$

from which the result (1.5) for $I_{n,0}(1)$ is easily recovered. The same result was also derived by Anis and Lloyd [2], [3]. In fact, these authors were prior to Sparre Andersen [1] in proving (4.10) for the special case when $P\{S_n \geqq 0\} = \frac{1}{2}$.

The probability density (4.9) remains the same when all inequalities $\geqq$ are replaced by strict inequalities $>$. We now employ a combinatorial result due to Spitzer [17, Thm. 2.1], rephrased as follows for the present purpose: "Let $y = (y_1, y_2, \cdots, y_n)$ be a vector such that $y_1 + y_2 + \cdots + y_n = 0$, but no other partial sum of distinct components vanishes. Let $y_{m+n} = y_m$, and $y(m) = (y_m, y_{m+1}, \cdots, y_{m+n})$, $m = 1, 2, \cdots, n$. Then exactly one of the n cyclic permutations $y(m)$ of y has the property that its successive partial sums are all positive except the last one which vanishes." Then it is easily seen that

(4.12) $$I_{n-1,0}(2) = (\pi^{1/2}/n) p\{S_n = 0\} = 1/n^{3/2},$$

since $p\{S_n = t\} = (\pi n)^{-1/2} \exp[-t^2/n]$.

4.2. $J_{n,0}(2)$, $J_{n,0}(1)$. Consider first the integral $J_{n,0}(2)$, as defined by (1.2). Proceeding as in § 4.1, we now introduce the new variables

(4.13) $$y_1 = x_1; \quad y_m = x_m - x_{m-1}, \qquad m = 2, 3, \cdots, n-1; \quad y_n = -x_n - x_{n-1}.$$

Then the integral $J_{n,0}(2)$ reduces to a form which is readily interpreted as a probability, namely,

(4.14) $$J_{n,0}(2) = P\{S_1 \geqq 0, \cdots, S_{n-1} \geqq 0, S_n \leqq 0\}$$

with S_m as defined in § 4.1. Compare (4.14) with (4.8), then it is obvious that

$$J_{n,0}(2)=I_{n-1,0}(1)-I_{n,0}(1)=-\frac{(-1/2)_n}{n!}, \tag{4.15}$$

in accordance with (1.6).

Secondly, the integral $J_{n+1,0}(1)$, as defined by (1.2), may be expressed in the form

$$\begin{aligned} J_{n+1,0}(1)=\pi^{-(n+1)/2}\int\overset{\cdot\cdot\cdot}{_{E^{n+1}}}\int \exp\Big[-x_1^2-2\sum_{m=2}^{n}x_m^2+2\sum_{m=1}^{n-1}x_mx_{m+1} \\ -2x_nx_{n+1}-x_{n+1}^2\Big]\prod_{m=1}^{n+1}H(x_m)\,dx_1\cdots dx_{n+1}, \end{aligned} \tag{4.16}$$

where E^{n+1} is the $(n+1)$-dimensional Euclidean space, and $H(x)$ stands for the unit step function, i.e., $H(x)=1$ for $x>0$ and $H(x)=0$ for $x<0$. Consider the exponent in (4.16) which is rewritten as

$$\begin{aligned} x_1^2+2\sum_{m=2}^{n}x_m^2-2\sum_{m=1}^{n-1}x_mx_{m+1}+2x_nx_{n+1}+x_{n+1}^2 \\ =\sum_{m=1}^{n-1}(x_m-x_{m+1})^2+(x_n+x_{n+1})^2. \end{aligned} \tag{4.17}$$

We now introduce the new variables

$$y_1=-x_n-x_{n+1};\quad y_m=x_{n+2-m}-x_{n+1-m},\quad m=2,3,\cdots,n;\quad y_{n+1}=-x_{n+1}. \tag{4.18}$$

Then, conversely,

$$x_{n+1-m}=y_{n+1}-\sum_{j=1}^{m}y_j,\qquad m=1,2,\cdots,n;\quad x_{n+1}=-y_{n+1}, \tag{4.19}$$

and (4.16) transforms into

$$\begin{aligned} J_{n+1,0}(1)=\pi^{-(n+1)/2}\int\overset{\cdot\cdot\cdot}{_{E^{n+1}}}\int \exp\Big[-\sum_{m=1}^{n}y_m^2\Big]\prod_{m=1}^{n}H\Big(y_{n+1}-\sum_{j=1}^{m}y_j\Big) \\ \cdot H(-y_{n+1})\,dy_1\cdots dy_{n+1}. \end{aligned} \tag{4.20}$$

Here, the integration with respect to y_{n+1} can be carried out, thus leading to

$$J_{n+1,0}(1)=\pi^{-(n+1)/2}\int\overset{\cdot\cdot\cdot}{_{E^{n}}}\int \exp\Big[-\sum_{m=1}^{n}y_m^2\Big]g(y)\,dy_1\cdots dy_n, \tag{4.21}$$

where

$$g(y)=-\min\Big[0,\max_{1\leqq m\leqq n}\sum_{j=1}^{m}y_j\Big]. \tag{4.22}$$

For a probabilistic interpretation of (4.21), let $y_1, y_2, \cdots, y_n$ be independent random variables with a common normal density function $\pi^{-1/2}\exp(-y_m^2)$, $m=1,2,\cdots,n$. Then $J_{n+1,0}(1)$, as given by (4.21), is equal to the expectation $\pi^{-1/2}E(g(y))$. By means of the notations

$$S_m=\sum_{j=1}^{m}y_j,\quad T_m=\sum_{j=2}^{m+1}y_j,\quad a^+=\max[0,a], \tag{4.23}$$

we reduce (4.22) to

$$
\text{(4.24)}\qquad \begin{aligned} g(y) = -\min\left[0, \max_{1\le m\le n} S_m\right] &= \max\left[0, \max_{1\le m\le n} S_m\right] - \max_{1\le m\le n} S_m \\ &= \max_{1\le m\le n} S_m^+ - y_1 - \max_{1\le m\le n-1} T_m^+. \end{aligned}
$$

Inserting (4.24) into (4.21), we may set

$$
\text{(4.25)}\qquad J_{n+1,0}(1) = \pi^{-1/2} E\left(\max_{1\le m\le n} S_m^+\right) - \pi^{-1/2} E\left(\max_{1\le m\le n-1} S_m^+\right),
$$

since the random variables S_m and T_m have the same distribution. The present result can be further reduced by means of the relation

$$
\text{(4.26)}\qquad E\left(\max_{1\le m\le n} S_m^+\right) = \sum_{m=1}^{n} \frac{1}{m} E(S_m^+),
$$

quoted from Spitzer [17, p. 330], and originally due to Kac [13, Thm. 4.1]. Thus we obtain as our final result

$$
\text{(4.27)}\qquad J_{n+1,0}(1) = \frac{\pi^{-1/2}}{n} E(S_n^+) = \frac{1}{\pi n^{3/2}} \int_0^\infty t e^{-t^2/n}\, dt = \frac{1}{2\pi n^{1/2}}.
$$

Acknowledgment. The author is indebted to Professor W. Schaafsma, University of Groningen, for bringing reference [16] to his attention.

REFERENCES

[1] E. Sparre Andersen, *On the fluctuations of sums of random variables II*, Math. Scand., 2 (1954), pp. 195–223.

[2] A. A. Anis and E. H. Lloyd, *On the range of partial sums of a finite number of independent normal variates*, Biometrika, 40 (1953), pp. 35–42.

[3] A. A. Anis, *The variance of the maximum of partial sums of a finite number of independent normal variates*, Ibid., 42 (1955), pp. 96–101.

[4] ———, *On the moments of the maximum of partial sums of a finite number of independent normal variates*, Ibid., 43 (1956), pp. 79–84.

[5] C. Atkinson, *A Wiener–Hopf integral equation arising in some inference and queueing problems*, Ibid., 61 (1974), pp. 277–283.

[6] J. Boersma, *Ray-optical analysis of reflection in an open-ended parallel-plane waveguide. I: TM case*, SIAM J. Appl. Math., 29 (1975), pp. 164–195.

[7] ———, *Ray-optical analysis of reflection in an open-ended parallel-plane waveguide: II—TE case*, Proc. IEEE, 62 (1974), pp. 1475–1481.

[8] ———, *Ray-optical analysis of reflection in an open-ended parallel-plane waveguide. III: Exterior radiation pattern*, to appear.

[9] A. Erdélyi, *Asymptotic Expansions*, Dover, New York, 1956.

[10] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. II, John Wiley, New York, 1966.

[11] A. Ghizzetti and A. Ossicini, *Studio di una particolare equazione integrale singolare di Wiener–Hopf*, Rend. Mat., 4 (1971), pp. 813–853.

[12] N. L. Johnson and S. Kotz, *Distributions in Statistics: Continuous Multivariate Distributions*, John Wiley, New York, 1972.

[13] M. Kac, *Toeplitz matrices, translation kernels and a related problem in probability theory*, Duke Math. J., 21 (1954), pp. 501–509.

[14] S. W. Lee and J. Boersma, *Ray-optical analysis of fields on shadow boundaries of two parallel plates*, J. Math. Phys., 16 (1975), pp. 1746–1764.

[15] B. NOBLE, *Methods Based on the Wiener–Hopf Technique for the Solution of Partial Differential Equations*, Pergamon Press, London, 1958.

[16] H.RUBEN, *The moments of the order statistics and of the range in samples from normal populations*, Contributions to Order Statistics, A. E. Sarhan and B. G. Greenberg, eds., John Wiley, New York, 1962, pp. 165–190.

[17] F. SPITZER, *A combinatorial lemma and its application to probability theory*, Trans. Amer. Math. Soc., 82 (1956), pp. 323–339.

[18] K. STEWARTSON, *On an integral equation*, Mathematika, 15 (1968), pp. 22–29.

SIAM J. MATH. ANAL.
Vol. 9, No. 3, June 1978

Copyright © 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0903-0001$01.00/0

AN INVERSE EIGENVALUE PROBLEM OF ORDER FOUR—AN INFINITE CASE*

JOYCE R. McLAUGHLIN†

Abstract. In this paper coefficients $A(s) \in C^3[0, 1]$ and $B(s) \in C^1[0, 1]$ are constructed so that two given positive sequences $\lambda_1 < \lambda_2 < \cdots$, and $\rho_1, \rho_2, \cdots$ are the eigenvalues and the corresponding normalization constants for the fourth order, self-adjoint, eigenvalue problem $y^{(4)} + (Ay^{(1)})^{(1)} + By - \lambda y = 0$, $y(0) = y^{(1)}(0) = y(1) = y^{(1)}(1) = 0$, $y^{(2)}(0) = 1$.

1. Introduction. The inverse eigenvalue problem to be considered here is that of assuming that two sequences of positive numbers $\lambda_1 < \lambda_2 < \cdots$ and $\rho_1, \rho_2, \cdots$ are given and then seeking to determine coefficients $A(s)$ and $B(s)$ and real numbers $M_{ij}, N_{ij}, i, j = 1, 2, 3, 4$ so that the set of real numbers $\lambda_1 < \lambda_2 < \cdots$ is the entire set of eigenvalues for the eigenvalue problem

$$y^{(4)}(s) + (A(s)y^{(1)}(s))^{(1)} + B(s)y(s) - \lambda y(s) = 0,$$

$$\sum_{j=1}^{4} [M_{ij}y^{(i-1)}(0) + N_{ij}y^{(i-1)}(1)] = 0, \qquad i = 1, 2, 3, 4,$$

and the sequence $\rho_1, \rho_2, \cdots$ is the corresponding set of normalization constants. (The normalization constants are squares of the L^2 norms of the eigenfunctions corresponding to the sequence of eigenvalues.) The method of solution is illustrated by assuming that the two sequences $\lambda_1 < \lambda_2 < \cdots$ and $\rho_1, \rho_2, \cdots$ have particular asymptotic forms. Then coefficients $A(s)$ and $B(s)$ are found so that the set $\lambda_1 < \lambda_2 < \cdots$ is the entire set of eigenvalues, and the set $\rho_1, \rho_2, \cdots$ is the corresponding set of normalization constants for the eigenvalue problem

$$\begin{aligned} &y^{(4)}(s) + [A(s)y^{(1)}(s)]^{(1)} + B(s)y(s) - \lambda y(s) = 0, \\ &y(0) = y^{(1)}(0) = y(1) = y^{(1)}(1) = 0, \qquad y^{(2)}(0) = 1. \end{aligned} \tag{1}$$

Requiring other asymptotic forms for $\lambda_1 < \lambda_2 < \cdots$ and $\rho_1, \rho_2, \cdots$ would result in different coefficients and other boundary conditions in the eigenvalue problem.

Interest in fourth order eigenvalue problems is fairly recent. However, inverse second order problems have been considered by a number of authors. Extensive work has been done by Borg [3], Marcenko [13], Krein [6], [7], Levinson [9], and Gel'fand and Levitan [4]. Roughly speaking in each of their papers sequences of eigenvalues and possibly a sequence of normalization constants are given and then a function $q(s)$ and corresponding boundary conditions are sought so that the given sequences are the eigenvalues and normalization constants for second order eigenvalue problems whose differential equation is $y^{(2)} + (\lambda - q(s))y = 0$. In [3], [13], [6], [7], [9], knowledge of two alternating sequences of eigenvalues is assumed and one function $q(s)$ and two sets of boundary conditions are found (to correspond with the sets of eigenvalues). In the

* Received by the editors August 14, 1975, and in revised form August 31, 1976.

† Department of Mathematical Sciences, Rennselaer Polytechnic Institute, Troy, New York 12181. This work was supported by National Science Foundation under Grant MPS75-08328.

paper by Gel'fand and Levitan [4] knowledge of a sequence of eigenvalues and a corresponding sequence of normalization constants is assumed and then the coefficient $q(s)$ plus one set of boundary conditions is determined to correspond to the sequence of eigenvalues and the sequence normalization constants. Furthermore, Levitan [11], [12] has shown that given the two sequences of eigenvalues in [3], [13], [6], [7], [9], the normalization constants associated with either sequence i.e., the spectral function, may be constructed. Hence if either two alternating sequences of eigenvalues are known or if the spectral function is known, the approach given by I. M. Gel'fand and B. M. Levitan may be used to find the unknown coefficient, $q(s)$, and boundary conditions in the second order inverse problem.

Some work in the fourth order inverse problem has been done by by Barcilon [1], [2], McKenna [14], and the author [15]. The work of V. Barcilon follows the approach of M. Krein. He shows that uniqueness of coefficients can be obtained for a fourth order problem from the knowledge of three, distinct, interlacing sequences of eigenvalues and three corresponding sets of boundary conditions. Further, he develops a method to find the coefficients in the fourth order differential equation when it is *known* that the three, interlacing, distinct sequences of eigenvalues are eigenvalues for eigenvalue problems which contain the given, corresponding sets of boundary conditions.

The present paper is an extension of the work in [15] which in turn was a generalization of the work of Gel'fand and Levitan [4]. In [15], $2n$ positive numbers were given and coefficients were found so that the positive numbers given were the *first n* eigenvalues and corresponding *first n* normalization constants for a fourth order eigenvalue problem. The remaining eigenvalues and normalization constants were chosen judiciously so that certain boundary conditions could be achieved for the resultant eigenvalue problem. The coefficients were determined as a finite sum of other functions, which in turn could be calculated merely by solving a finite set of nonhomogeneous linear equations.

More specifically, the solution in [15] was found in the following way. We assumed that $2n$ positive numbers $\lambda_1 < \lambda_2 < \cdots < \lambda_n, \rho_1, \rho_2, \cdots, \rho_n$ were given. Then, we let

$$Z_\lambda = \beta \frac{[\sin \lambda^{1/4} s - \sinh \lambda^{1/4} s]}{2\lambda^{1/2}} + \frac{[\cosh \lambda^{1/4} s - \cos \lambda^{1/4} s]}{2\lambda^{1/2}},$$

where

$$\beta = \frac{\cosh \lambda^{1/4} - \cos \lambda^{1/4}}{\sinh \lambda^{1/4} - \sin \lambda^{1/4}}.$$

(This definition for Z_λ is slightly different than that in [15] but it does not alter the results.)

Further, let $\lambda_1^* < \lambda_2^* < \cdots$ be the eigenvalues for the eigenvalue problem $Z^{(4)} + \lambda Z = 0$, $Z(0) = Z^{(1)}(0) = Z(1) = Z^{(1)}(1) = 0$. Then $Z_{\lambda_i^*}$ is the eigenfunction corresponding to λ_i^* with $Z^{(2)}_{\lambda_i^*}(0) = 1$, $i = 1, 2, \cdots$. Define $\rho_i^* = \int_0^1 [Z_{\lambda_i^*}(s)]^2 \, ds$, $i = 1, 2, \cdots$, and

$$f_n(s, t) = \sum_{i=1}^n \left[\frac{Z_{\lambda_i}(s) Z_{\lambda_i}(t)}{\rho_i} - \frac{Z_{\lambda_i^*}(s) Z_{\lambda_i^*}(t)}{\rho_i^*} \right].$$

With these definitions it was shown that *the* solution $K_n(s,t)$ (and all of its derivatives with respect to s and t) of $f_n(s,t)+\int_0^s f_n(t,y)K_n(s,y)\,dy+K_n(s,t)=0$ could be determined from solutions of linear nonhomogeneous equations. And, in addition, it was shown that if $A_n(s)=-4(d/ds)K_n(s,s)$, and $B_n(s)=-A_nK_{n,s}|_{t=s}+2(K_{n,ss}-K_{n,tt})_t|_{t=s}-2(d^3/ds^3)K_n(s,s)$, and $y_i^n(s)=Z_{\lambda_i}(s)+\int_0^s K_n(s,t)Z_{\lambda_i}(t)\,dt$, then $\lambda_1<\lambda_2<\cdots<\lambda_n$ would be the first n eigenvalues, with corresponding eigenfunctions $y_i^n(s)$, $i=1,2,\cdots,n$, for the problem

$$y^{(4)}+(A_ny^{(1)})^{(1)}+B_ny-\lambda y=0,$$
$$y(0)=y^{(1)}(0)=y(1)=y^{(1)}(1)=0,$$

with $\rho_i=\int_0^1[y_i^n(s)]^2\,ds$, $i=1,2,\cdots,n$. The nature of the construction of the above coefficients and eigenfunctions also yielded that the remaining eigenvalues for the above problem would be λ_i^*, $i=n+1,n+2,\cdots$, with associated eigenfunction $y_i^n=Z_{\lambda_i^*}(s)+\int_0^s K_n(s,t)Z_{\lambda_i^*}(t)\,dt$, and normalization constant $\rho_i^*=\int_0^1[y_i^n(s)]^2\,ds$, $i=n+1,n+2,\cdots$.

We seek now to show that if the infinite set of numbers $\lambda_1<\lambda_2<\cdots$ and $\rho_1,\rho_2,\cdots$ are given positive numbers that satisfy appropriate asymptotic forms, then there exists $A(s)\in C^1[0,1]$ and $B(s)\in C[0,1]$ such that $\lambda_1<\lambda_2<\cdots$ are all the eigenvalues for the eigenvalue problem

$$y^{(4)}+(Ay^{(1)})^{(1)}+By-\lambda y=0,$$
$$y(0)=y^{(1)}(0)=y(1)=y^{(1)}(1)=0.$$

And if $y_i(s)$ is the eigenfunction associated with λ_i, with $y_i^{(2)}(0)=1$, then $\rho_i=\int_0^1[y_i(s)]^2\,ds$, $i=1,2,\cdots$. In particular, we seek to show that if $\{\lambda_i\}_{i=1}^\infty$ and $\{\rho_i\}_{i=1}^\infty$ satisfy the appropriate asymptotic forms then $A(s)$, $B(s)$ and $y_i(s)$, $i=1,2,\cdots$, can be determined by

$$A(s)=\lim_{n\to\infty}A_n(s),\quad B(s)=\lim_{n\to\infty}B_n(s),\quad y_i(s)=\lim_{n\to\infty}y_i^n(s),$$

where $A_n(s)$, $B_n(s)$ and $y_i^n(s)$, $i,n=1,2,\cdots$, are as given above. This will give us sufficient conditions for convergence if we seek to obtain solutions close to the "real" solution by applying the technique of [15]. Furthermore, it should be noted that it is not necessary to know, *in advance*, that the sequence of eigenvalues and the sequence of normalization constants arise from some unknown fourth order eigenvalue problem; that is, the *existence* of coefficients $A(s)$ and $B(s)$ so that $\lambda_1<\lambda_2<\cdots$ and $\rho_1,\rho_2,\cdots$ are eigenvalues and normalization constants for the fourth order eigenvalue problem is contained in the proofs in this paper.

The procedure for proving the above is as follows. In § 2, we require the sets $\{\lambda_i\}_{i=1}^\infty$ and $\{\rho_i\}_{i=1}^\infty$ to have appropriate asymptotic forms so that $f_n(s,t)$ and all derivatives, up to and including the order four, converge uniformly as $n\to\infty$ to

$$f(s,t)=\sum_{i=1}^\infty\left[\frac{Z_{\lambda_i}(s)Z_{\lambda_i}(t)}{\rho_i}-\frac{Z_{\lambda_i^*}(s)Z_{\lambda_i^*}(t)}{\rho_i^*}\right]$$

and the corresponding derivatives of $f(s,t)$. Then *the* solution, $K(s,t)$ of the integral equation $f(s,t)+\int_0^s f(t,y)K(s,y)\,dy+K(s,t)=0$ has continuous derivatives up to and including the order four. Furthermore,

$$\lim_{n\to\infty}\frac{\partial^{i+j}}{\partial s^i\,\partial t^j}K_n(s,t)=\frac{\partial^{i+j}}{\partial s^i\,\partial t^j}K(s,t),\qquad i,j=0,1,2,3,4,\quad 0\leqq i+j\leqq 4.$$

Once this is determined, it can be shown that there exists $A(s)\in C^3[0,1]$ and $B(s)\in C^1[0,1]$, $y_i(s)\in C^4[0,1]$ such that $A(s)=\lim_{n\to\infty}A_n(s)$, $B(s)=\lim_{n\to\infty}B_n(s)$, $y_i(s)=Z_{\lambda_i}(s)+\int_0^s K(s+t)Z_{\lambda_i}(t)\,dt$, and $y_i(s)$ satisfies the differential equation

$$y_i^{(4)}+(Ay_i^{(1)})^{(1)}+By_i-\lambda_i y_i=0,\qquad i=1,2,\cdots.$$

In § 3, it is shown that $\rho_i=\int_0^1[y_i(s)]^2\,ds$ and that each $y_i(s)$ satisfies the boundary conditions of (1), $i=1,2,\cdots$. Further the set $\lambda_1<\lambda_2<\cdots$ is shown to be the entire sequence of eigenvalues for the eigenvalue problem (1). Finally in § 4 more general inverse problems are discussed. That is, application of the techniques of §§ 2 and 3 are discussed for the cases when different asymptotic forms are known and/or other boundary conditions are desired.

2. Convergence of A_n, B_n, y_i^n, as $n\to\infty$. In this section we will determine conditions on the positive numbers $\lambda_1<\lambda_2<\cdots$ and $\rho_1,\rho_2,\cdots$ in order that the finite sums

$$\frac{\partial^{j+k}}{\partial s^j\,\partial t^k}f_n(s,t)=\frac{\partial^{i+k}}{\partial s^j\,\partial t^k}\sum_{i=1}^{n}\frac{Z_{\lambda_i}(s)Z_{\lambda_i}(t)}{\rho_i}-\frac{Z_{\lambda_i^*}(s)Z_{\lambda_i^*}(t)}{\rho_i^*}$$

converge uniformly to

$$\frac{\partial^{j+k}}{\partial s^j\,\partial t^k}f(s,t)=\frac{\partial^{j+k}}{\partial s^j\,\partial t^k}\sum_{i=1}^{\infty}\frac{Z_{\lambda_i}(s)Z_{\lambda_i}(t)}{\rho_i}-\frac{Z_{\lambda_i^*}(s)Z_{\lambda_i^*}(t)}{\rho_i^*}$$

for $j,k=0,1,2,3,4$, $0\leqq j+k\leqq 4$. We will also determine conditions on the positive sequence $\lambda_1<\lambda_2<\cdots$ in order for the set of functions $\{Z_{\lambda_i}(s)\}_{i=1}^{\infty}$ to be a complete set of functions in $L^2[0,1]$. Using these results it will be shown that there is a unique solution $K(s,t)\in C[0\leqq t\leqq s\leqq 1]$ of

$$f(s,t)+\int_0^s f(t,y)K(s,y)\,dy+K(s,t)=0 \tag{2}$$

and that $K(s,t)$ has derivatives with respect to s and t, for $0\leqq t\leqq s\leqq 1$, up to and including the order 4 with the property that

$$\lim_{n\to\infty}\frac{\partial^{j+k}}{\partial s^j\,\partial t^k}K_n(s,t)=\frac{\partial^{j+k}}{\partial s^j\,\partial t^k}K(s,t),$$

uniformly for $0\leqq t\leqq s\leqq 1$, j, $k=0,1,2,3,4$, $0\leqq j+k\leqq 4$ where $K_n(s,t)$ is the unique solution of $f_n(s,t)+\int_0^s f_n(t,y)K_n(s,y)\,dy+K_n(s,t)=0$; see [15].

Further, suppose $A_n(s)$ and $B_n(s)$ are the coefficients, calculated as in [15], so that the fourth order eigenvalue problem

$$y^{(4)}+(A_ny^{(1)})^{(1)}+B_ny-\lambda y=0,$$
$$y(0)=y^{(1)}(0)=y(1)=y^{(1)}(1)=0$$

will have $\lambda_1<\lambda_2<\cdots<\lambda_n<\lambda^*_{n+1}<\lambda^*_{n+2}<\cdots$ as *all* of its eigenvalues, and so that if y^n_i is the eigenfunction associated with the ith eigenvalue, with $(y^n_i)^{(2)}(0)=1$, then

$$\int_0^1 [y^n_i]^2\,ds=\begin{cases}\rho_i, & i=1,2,\cdots,n,\\ \rho^*_i, & i=n+1,n+2,\cdots.\end{cases}$$

Then it will be shown that there exist functions $A(s)\in C^3[0,1]$, $B(s)\in C^1[0,1]$, $y_i(s)\in C^4[0,1]$, $i=1,2,\cdots$, such that

$$\lim_{n\to\infty}[A_n(s)]^{(k)}=[A(s)]^{(k)}\quad\text{uniformly for } 0\leqq s\leqq 1,\quad k=0,1,$$

$$\lim_{n\to\infty}B_n(s)=B(s)\quad\text{uniformly for } 0\leqq s\leqq 1,$$

$$\lim_{n\to\infty}[y^n_i(s)]^{(k)}=[y_i(s)]^{(k)}\quad\text{uniformly for } 0\leqq s\leqq 1,$$
$$k=0,1,2,3,4,\quad i=1,2,3,\cdots$$

and such that $y^{(4)}_i+(Ay^{(1)}_i)^{(1)}+By_i-\lambda_iy_i=0$. (It should be noted that the assumption that $\lambda_n<\lambda^*_{n+1}$ is achieved for large n through the asymptotic form required of the sequence $\{\lambda_i\}_{i=1}^{\infty}$.)

We begin with the following five theorems and two corollaries.

THEOREM 1. *Let $\lambda_1<\lambda_2<\cdots$ and $\rho_1,\rho_2,\cdots$ be given sequences of positive numbers. Let Z_λ, λ^*_i, ρ^*_i, $i=1,2,\cdots$, be given as in the Introduction. Suppose that $(\lambda_i)^{1/4}=(\lambda^*_i)^{1/4}+P_i$, $1/(\lambda_i\rho_i)=1/(\lambda^*_i\rho^*_i)+R_i$ and that the two sums $\sum_{i=1}^{\infty}(\lambda_i)^{k/4}|P_i|$ and $\sum_{i=1}^{\infty}(\lambda_i)^{k/4}|R_i|$ converge for $k=0,1,2,3,4$. If we define*

$$f(s,t)=\sum_{i=1}^{\infty}\frac{Z_{\lambda_i}(s)Z_{\lambda_i}(t)}{\rho_i}-\frac{Z_{\lambda^*_i}(s)Z_{\lambda^*_i}(t)}{\rho^*_i},$$

then the series

$$\sum_{i=1}^{\infty}\left\{\frac{\left[\frac{d^j}{ds^j}Z_{\lambda_i}(s)\right]\left[\frac{d^k}{dt^k}Z_{\lambda_i}(t)\right]}{\rho_i}-\frac{\left[\frac{d^j}{ds^j}Z_{\lambda^*_i}(s)\right]\left[\frac{d^k}{dt^k}Z_{\lambda^*_i}(t)\right]}{\rho^*_i}\right\}$$

convergences uniformly to continuous functions $(\partial^{j+k}/\partial s^j\,\partial t^k)f(s,t)$ for $j,k=0,1,2,3,4$ and $0\leqq j+k\leqq 4$, $0\leqq t\leqq s\leqq 1$.

Proof. The proof consists of a careful examination of the term

$$\frac{Z_{\lambda_i}(s)Z_{\lambda_i}(t)}{\rho_i}-\frac{Z_{\lambda^*_i}(s)Z_{\lambda^*_i}(t)}{\rho^*_i}$$

and its derivatives. For ease of notation, we let $\lambda=\mu^4$, for λ and μ positive. Then after suitable rearrangement of terms, $Z_\lambda(s)=(1/(2\mu^2))[J(\mu,s)+e^{-\mu}H(\mu,s)]$

where $J(\mu, s) = \sin \mu s - \cos \mu s + e^{-\mu s} + e^{-\mu+\mu s}(\cos \mu - \sin \mu)$ and where $H(\mu, s)$ has the property that there exists $M > 0$, independent of μ and s, such that $|(\partial^k/\partial s^k)H(\mu, s)| \leqq M\mu^k$ for $k = 0, 1, 2, 3, 4$ and $\mu \geqq 0$. Further, using the above estimates, we can write

$$4\lambda_i^*\rho_i^* = 4\lambda_i^* \int_0^1 [Z_{\lambda_i^*}(s)]^2 \, ds = \int_0^1 [J(\mu_i^*, s)]^2 \, ds + e^{-\mu_i^*} I_1(\mu_i^*)$$

where $(\mu_i^*)^4 = \lambda_i^*$ and $I_1(\mu_i^*)$ is uniformly bounded for $i = 1, 2, \cdots$. Straightforward integration then yields that $4\lambda_i^*\rho_i^* = 1 + (1/\mu_i^*) \cdot (\cos \mu_i^* - \sin \mu_i^*) \cos \mu_i^* + e^{-\mu_i^*} I_2(\mu_i^*)$ where $I_2(\mu_i^*)$ is uniformly bounded for $i = 1, 2, \cdots$. We then can write

$$\begin{aligned}
&\frac{Z_{\lambda_i}(s)Z_{\lambda_i}(t)}{\rho_i} - \frac{Z_{\lambda_i^*}(s)Z_{\lambda_i^*}(t)}{\rho_i^*} \\
&\quad = \frac{1}{4\lambda_i\rho_i}\{Q(\mu_i, s, t) + J(\mu_i, s)\, e^{-\mu_i} H(\mu_i, t) \\
&\qquad\quad + J(\mu_i, t)\, e^{-\mu_i} H(\mu_i, s) + e^{-2\mu_i} H(\mu_i, t) H(\mu_i, s)\} \\
&\quad\quad - \frac{1}{4\lambda_i^*\rho_i^*}\{Q(\mu_i^*, s, t) + J(\mu_i^*, s)\, e^{-\mu_i^*} H(\mu_i^*, t) \\
&\qquad\quad + J(\mu_i^*, t)\, e^{-\mu_i^*} H(\mu_i^*, s) + e^{-2\mu_i^*} H(\mu_i^*, t) H(\mu_i^*, s)\}
\end{aligned}$$

where

$$\begin{aligned}
Q(\mu, s, t) = {} & \cos \mu(s-t) - \sin \mu(s+t) + e^{-\mu s}(\sin \mu t - \cos \mu t) \\
& + e^{-\mu+\mu s}(-\cos \mu(1-t) + \sin \mu(1+t) + e^{-\mu t}(\sin \mu s - \cos \mu s) \\
& + e^{-\mu(t+s)} + e^{-\mu+\mu(-t+s)}(\cos \mu - \sin \mu) \\
& + e^{-\mu+\mu t}(-\cos \mu s + \sin \mu s) + e^{-\mu+\mu(t-s)}(\cos \mu - \sin \mu) \\
& + e^{-2\mu+\mu(t+s)}(1 - \sin 2\mu).
\end{aligned}$$

It can be shown that the set $\{\mu_i^*\}_{i=1}^\infty$ is the set of zeros of the function $1 - \cos \mu \cosh \mu$. Hence it can be shown that $\lim_{i\to\infty} (\mu_i^* - i\pi - \pi/2) = 0$. In addition, by hypothesis we have $\lim_{i\to\infty} (\mu_i^* - \mu_i) = 0$; hence $\lim_{i\to\infty} (\mu_i - i\pi - \pi/2) = 0$. Therefore

$$\begin{aligned}
\sum_{i=1}^\infty \frac{\partial^{k+j}}{\partial s^j \, \partial t^k}\bigg[\left(\frac{1}{4\lambda_i\rho_i}\right)\{J(\mu_i, s)\, e^{-\mu_i} H(\mu_i, t) + J(\mu_i, t)\, e^{-\mu_i} H(\mu_i, s) \\
+ e^{-2\mu_i} H(\mu_i, t) H(\mu_i, s)\}\bigg]
\end{aligned}$$

converges uniformly for $0 \leqq t \leqq s \leqq 1$, $k, j = 0, 1, 2, 3, 4$, $0 \leqq k + j \leqq 4$. A similar statement holds when μ_i is replaced by μ_i^* in the above expression. We need only concern ourselves, then, with the terms

$$\frac{Q(\mu_i, s, t)}{4\lambda_i\rho_i} - \frac{Q(\mu_i^*, s, t)}{4\lambda_i^*\rho_i^*}$$

Following the notation in the hypothesis of this theorem, we write

$$\frac{Q(\mu_i, s, t)}{4\lambda_i\rho_i} - \frac{Q(\mu_i^*, s, t)}{4\lambda_i^*\rho_i^*} = R_iQ(\mu_i, s, t) + \frac{1}{\rho_i^*4\lambda_i^*}[Q(\mu_i, s, t) - Q(\mu_i^*, s, t)].$$

It is clear from the definition of $Q(\mu, s, t)$ that there exists $M>0$ such that $|(\partial^{k+j}/\partial s^j\,\partial t^k)Q(\mu, s, t)| \leqq M\mu^{k+j}$, $k, j = 0, 1, 2, 3, 4$, $0 \leqq k+j \leqq 4$. Hence, we have that $\sum_{i=1}^{\infty}(\partial^{k+j}/\partial s^j\,\partial t^k)[R_iQ(\mu_i, s, t)]$ converges uniformly for $0 \leqq k+j \leqq 4$, $k, j = 0, 1, 2, 3, 4$. In order to show that

$$\sum_{i=1}^{\infty} \frac{1}{\rho_i^*4\lambda_i^*} \frac{\partial^{k+j}}{\partial s^j\,\partial t^k}[Q(\mu_i, s, t) - Q(\mu_i^*, s, t)]$$

converges uniformly, we consider first the difference of the first terms in $Q(\mu_i, s, t)$ and $Q(\mu_i^*, s, t)$. That is, we consider

$$\sum_{i=1}^{\infty} \frac{1}{\rho_i^*4\lambda_i^*} \frac{\partial^{k+j}}{\partial s^j\,\partial t^k}[\cos \mu_i(s-t) - \cos \mu_i^*(s-t)].$$

Straightforward computation yields that there exists $M>0$ such that

$$|(\partial^{k+j}/\partial s^j\,\partial t^k)[\cos \mu_i(s-t) - \cos \mu_i^*(s-t)]| \leqq M(\mu_i^*)^{k+j}|P_i|.$$

Hence

$$\sum_{i=1}^{\infty} \frac{1}{\rho_i^*4\lambda_i^*} \frac{\partial^{k+j}}{\partial s^j\,\partial t^k}[\cos \mu_i(s-t) - \cos \mu_i^*(s-t)]$$

converges uniformly for $k, j, k+j = 0, 1, 2, 3, 4$. Similar bounds yield the required uniform convergence of

$$\sum_{i=1}^{\infty} \frac{1}{\rho_i^*4\lambda_i^*} \frac{\partial^{k+j}}{\partial s^j\,\partial t^k}[Q(\mu_i, s, t) - Q(\mu_i^*, s, t)], \qquad k, j, k+j = 0, 1, 2, 3, 4.$$

The fact that the derivative $(\partial^{j+k}/\partial s^j\,\partial t^k)f(s, t)$ is continuous for $j, k = 0, 1, 2, 3, 4$, $0 \leqq j+k \leqq 4$, $0 \leqq t \leqq s \leqq 1$ follows from the uniform convergence of the series.

Remark. It should be noted that in differentiating the series

$$f(s, t) = \sum_{i=1}^{\infty}\left[\frac{Z_{\lambda_i}(s)Z_{\lambda_i}(t)}{\rho_i} - \frac{Z_{\lambda_i^*}(s)Z_{\lambda_i^*}(t)}{\rho_i^*}\right]$$

each successive differentiation introduces a factor of $(\lambda_i)^{1/4}$ into the numerator of the term obtained by differentiating $Z_{\lambda_i}(s)Z_{\lambda_i}(t)/\rho_i$ and introduces a factor of $(\lambda_i^*)^{1/4}$ into the numerator of the term obtained by differentiating $Z_{\lambda_i^*}(s)Z_{\lambda_i^*}(t)/\rho_i^*$. Hence it can be seen from the proof of Theorem 1 that in order for

$$f(s, t) = \sum_{i=1}^{\infty}\left[\frac{Z_{\lambda_i}(s)Z_{\lambda_i}(t)}{\rho_i} - \frac{Z_{\lambda_i^*}(s)Z_{\lambda_i^*}(t)}{\rho_i^*}\right]$$

to converge uniformly for $0 \leqq t \leqq s \leqq 1$ we need only have $\sum_{i=1}^{\infty}|P_i| < \infty$ and $\sum_{i=1}^{\infty}|R_i| < \infty$. Further, first order derivatives of $f(s, t)$ can be obtained by differentiating the series, representing $f(s, t)$, termwise, if $\sum_{i=1}^{\infty}\lambda_i^{1/4}|P_i| < \infty$ and

$\sum_{i=1}^{\infty} (\lambda_i^*)^{1/4}|R_i| < \infty$, and kth order derivatives can be obtained by differentiating the series k times, termwise, if $\sum_{i=1}^{\infty} (\lambda_i)^{k/4}|P_i| < \infty$ and $\sum_{i=1}^{\infty} (\lambda_i^*)^{k/4}|R_i| < \infty$, $k = 1, 2, \cdots$.

Before proving Theorem 2 we remark that we have changed notation in Theorem 2 to facilitate the notation in the proof. Accordingly, we have let μ be the positive root $\mu = \lambda^{1/4}$ for $\lambda > 0$ and we let $Z^{\mu}(s) = 2\mu^2(\sinh \mu - \sin \mu)Z_\lambda(s)$. Since we are concerned here with the completeness of the functions $\{Z_{\lambda_i}(s)\}_{i=1}^{\infty}$ these changes do not alter our completeness result. To aid the reader Corollary 2 is written out in terms of $\{Z_{\lambda_i}(s)\}_{i=1}^{\infty}$.

THEOREM 2. *Let* $Z^{\mu}(s) = (\cosh \mu - \cos \mu)(\sin \mu s - \sinh \mu s) + (\sinh \mu - \sin \mu)(\cosh \mu s - \cos \mu s)$, *and let* $0 < \mu_1 < \mu_2 < \cdots$ *be a given sequence of positive numbers. Let* $\mu_0 = 0$. *Let* $\Lambda(u)$ *be the number of* $\mu_n \leqq u$, $n = 0, 1, 2, \cdots$. *Then if for* $v > 1$, *and some* $b > 1$,

$$\int_1^v \frac{\Lambda(u)}{u}\, du > \frac{v}{\pi} - \frac{1}{4b} \log v - C$$

for some constant C, then the set $\{Z^{\mu_i}\}_{i=1}^{\infty}$ *is complete on* $L^2(0, 1)$.

Remark. The definition of completeness of $\{Z^{\mu_i}(s)\}_{i=1}^{\infty}$ on $L^2(0, 1)$ used here is that if $f \in L^2(0, 1)$ and $\int_0^1 Z^{\mu_i}(s)f(s)\, ds = 0$, $i = 1, 2, \cdots$, implies that $f = 0$, a.e. on $0 \leqq s \leqq 1$, then $\{Z^{\mu_i}\}_{i=1}^{\infty}$ is complete on $L^2(0, 1)$.

Proof. The proof is by contradiction. Assume, contrary to what we want to prove that there exist $f(s) \in L^2(0, 1)$ such that $f(s) \neq 0$ on a set of positive measure and such that the analytic function

$$H(\mu) = \int_0^1 f(s)Z^{\mu}(s)\, ds$$

has the property that $H(\mu_n) = 0$, $n = 1, 2, \cdots$. Further $H[(\lambda_n^*)^{1/4}]$ cannot be zero for every $n = 1, 2, \cdots$, since $f(s) \neq 0$ on a set of positive measure and $\{Z^{(\lambda_n^*)^{1/4}}\}_{n=1}^{\infty}$ is a complete set on $L^2(0, 1)$. Hence $H(\mu) \not\equiv 0$. In addition we observe that $H(-\mu_n) = H(i\mu_n) = H(-i\mu_n) = 0$, $n = 1, 2, \cdots$. Also since $Z^{\mu}(s)$ has a fifth order zero at $\mu = 0$, then $H(\mu)$ has a fifth order zero at $\mu = 0$.

Denote $n(r)$ as the number of zeros of $H(\mu)$ when $|\mu| \leqq r$. We now obtain an upper bound for $\int_1^r n(u)/u\, du$. By Jensen's theorem, see e.g. [10, p. 243],

$$\int_1^r \frac{n(u)}{u}\, du \leqq \frac{1}{2\pi} \int_0^{2\pi} \log |H(re^{i\theta})|\, d\theta + a_1$$

where a_1 is independent of r. From the form of $Z^{\mu}(s)$, it is seen that when $\mu = re^{i\theta}$, we have the estimate

$$\begin{aligned} |Z^{\mu}(s)| &\leqq d \max \left[e^{r|\cos\theta|} e^{rs|\sin\theta|}, e^{r(1-s)|\cos\theta|}, e^{r(1-s)|\sin\theta|}, e^{rs|\cos\theta|} e^{r|\sin\theta|}\right] \\ &\leqq d\, e^{r|\cos\theta|} e^{r|\sin\theta|} \end{aligned}$$

where d is a constant independent of r, s and θ. Substituting this estimate in the above inequality, and observing that $\int_0^{2\pi} [|\cos\theta| + |\sin\theta|]\, d\theta = 8$, we can obtain

$$\int_1^r \frac{n(u)}{u}\, du \leqq \frac{4r}{\pi} + a_2$$

where a_2 is independent of r.

We will now achieve a contradiction in the following way. We construct the analytic function

$$F(\mu)=\mu^4\prod_{n=1}^{\infty}\left(1-\frac{\mu^4}{\lambda_n^4}\right),$$

see e.g. [5, Chap. 1], which has a fourth order zero at $\mu=0$, and simple zeros at $\mu=\mu_n, -\mu_n, i\mu_n, -i\mu_n, n=1,2,\cdots$. Then the function $\phi(\mu)=H(\mu)/F(\mu)$ is also analytic for all finite μ. The number of zeros, $n_1(r)$, of $\phi(\mu)$ for $|\mu|\leqq r$ is $n_1(r)=n(r)-4\Lambda(r)$ and hence for $r>1$

$$\int_1^r \frac{n_1(r)}{r}\,dr<\frac{4r}{\pi}+a_2-\left[\frac{4r}{\pi}-\frac{1}{b}\log r-C\right]=\frac{1}{b}\log r+a_4$$

where a_4 is independent of r. The coefficient of $\log r$ is less than 1, and $n_1(r)$ is integer valued, nonnegative, and nondecreasing. Hence $n_1(r)\equiv 0$ and $\phi(\mu)$ has no zeros in the finite plane. But it is known that $H(\mu)$ has a fifth order zero at $\mu=0$ while $F(\mu)$ has exactly a fourth order zero at $\mu=0$. Hence, $\phi(0)=0$. This contradiction proves the result.

COROLLARY 1. *If $\mu_n>0$, $n=1,2,\cdots$, and $\mu_n=((2n+1)/2)\pi+O(1/n)$, $n=1,2,\cdots$, then $\{Z^{\mu_n}(s)\}_{n=1}^{\infty}$ is a complete set on $L^2(0,1)$.*

Proof. The result follows easily once it is observed that for large μ, there exists $k>1$ such that $\Lambda(\mu)\geqq\mu/\pi-1/(4k)$.

COROLLARY 2. *If $\lambda_1<\lambda_2<\cdots$ and $\lambda_i^{1/4}=((2i+1)/2)\pi+O(1/i)$, $i=1,2,\cdots$, then $\{Z_{\lambda_i}(s)\}_{i=1}^{\infty}$ is a complete set on $L^2(0,1)$.*

Proof. Since Z^{μ_i} is a constant multiple of Z_{λ_i}, we have that $\{Z_{\lambda_i}\}_{i=1}^{\infty}$ is a complete set on $L^2(0,1)$ iff $\{Z^{\mu_i}\}_{i=1}^{\infty}$ is a complete set on $L^2(0,1)$. Hence, Corollary 2 is just a restatement of Corollary 1.

Remark. The hypothesis in Corollary 2 is implied by the hypothesis of Theorem 1. This is seen by noting first that the sequence $\{\lambda_i^*\}_{i=1}^{\infty}$ is determined as the set of fourth powers of the positive roots of the equation

$$\cosh(\lambda^{1/4})\cos(\lambda^{1/4})-1=0.$$

Further, it is easily seen, by graph of the function $\cosh\lambda^{1/4}\cos\lambda^{1/4}$ that the sequence $\{\lambda_i^*\}_{i=1}^{\infty}$ has the property $\lim_{i\to\infty}[(\lambda_i^*)^{1/4}-((2i+1)/2)\pi]=0$. By letting $(\lambda_i^*)^{1/4}=((2i+1)/2)\pi+\alpha_i$, it can then be shown that $\lim_{i\to\infty}[\alpha_i e^{[((2i+1)/2)\pi]}-2]=0$. Hence we have that $(\lambda_i^*)^{1/4}=((2i+1)/2)\pi+O(1/i)$. In addition the hypothesis of Theorem 1 states that $(\lambda_i)^{1/4}=(\lambda_i^*)^{1/4}+P_i$ where $\sum_{i=1}^{\infty}|P_i|<\infty$. This yields that the sequence $\{P_i\}_{i=1}^{\infty}$ satisfies $P_i=O(1/i)$. Hence $(\lambda_i)^{1/4}=((2i+1)/2)\pi+O(1/i)$.

THEOREM 3. *Let $Z_{\lambda_i}(t)$, $Z_{\lambda_i^*}(t)$, ρ_i, ρ_i^* be defined as above. Assume that*

$$f(s,t)=\sum_{i=1}^{\infty}\left[\frac{Z_{\lambda_i}(s)Z_{\lambda_i}(t)}{\rho_i}-\frac{Z_{\lambda_i^*}(s)Z_{\lambda_i^*}(t)}{\rho_i^*}\right]$$

converges uniformly for $0\leqq t\leqq s\leqq 1$, and that $g(s,t)$ is continuous for $0\leqq t\leqq s\leqq 1$. Further assume that the set of functions $\{Z_{\lambda_i}(s)\}_{i=1}^{\infty}$ is complete on the interval $0\leqq s\leqq 1$.

Then there exists exactly one solution $k(s,t)$ continuous in t, $0 \leqq t \leqq s$, for each s of

$$g(s,t)+\int_0^s f(t,y)k(s,y)\,dy+k(s,t)=0.$$

Proof. The Fredholm theory tells us that for each s there exists exactly one solution, continuous in t of the above equation if the only continuous solution $h(t)$ of

$$\int_0^s f(t,y)h(y)\,dy+h(t)=0$$

is $h\equiv 0$. Hence, fix s and assume $h(t)$ is a continuous solution of $\int_0^s f(t,y)h(y)\,dy+h(t)=0$. Then multiply this equation by $h(t)$ and integrate from 0 to s to obtain

$$\begin{aligned}0=&\int_0^s [h(t)]^2\,dt\\&+\sum_{i=1}^{\infty}\left\{\frac{[\int_0^s h(t)Z_{\lambda_i}(t)\,dt][\int_0^s h(y)Z_{\lambda_i}(y)\,dy]}{\rho_i}\right.\\&\left.-\frac{[\int_0^s h(t)Z_{\lambda_i^*}(t)\,dt\int_0^s h(y)Z_{\lambda_i^*}(y)\,dy]}{\rho_i^*}\right\}.\end{aligned}$$

(The integrals may be taken inside the summation sign since the infinite sum which defines $f(s,t)$ converges uniformly.) Parseval's equation tells us that

$$\int_0^s [h(t)]^2\,dt=\sum_{i=1}^{\infty}\frac{[\int_0^s h(t)Z_{\lambda_i^*}(t)\,dt]^2}{\rho_i^*}$$

and hence we have that

$$\sum_{i=1}^{\infty}\frac{[\int_0^s h(t)Z_{\lambda_i}(t)\,dt]^2}{\rho_i}=0.$$

Since this is an infinite sum of nonnegative real numbers, it must now be true that

$$\int_0^s h(t)Z_{\lambda_i}(t)\,dt=0,\qquad i=1,2,\cdots.$$

But $\{Z_{\lambda_i}(t)\}_{i=1}^{\infty}$ is a complete set of functions on $0\leqq t\leqq 1$ and hence is a complete set of functions on any subinterval. Thus $h(t)\equiv 0$.

THEOREM 4. *Suppose that there exist functions $e_n(t,s)$, $n=1,2,\cdots$, $e(t,s)$ continuous on $0\leqq t\leqq s\leqq 1$, and functions $d_n(t,y,s)$, $n=1,2,\cdots$, $d(t,y,s)$ continuous for $0\leqq t$, $y\leqq s\leqq 1$ such that $\lim_{n\to\infty} e_n(t,s)=e(t,s)$ uniformly for $0\leqq t\leqq s\leqq 1$ and $\lim_{n\to\infty} d_n(t,y,s)=d(t,y,s)$ uniformly for $0\leqq t$, $y\leqq s\leqq 1$. Also, suppose there exists a unique continuous solution $Q_n(t,s)$, $0\leqq t\leqq s\leqq 1$, of the integral equation*

$$e_n(t,s)+\int_0^s d_n(t,y,s)Q_n(s,y)\,dy+Q_n(s,t)=0,\qquad n=1,2,\cdots,$$

and there exists a unique continuous solution $Q(s,t)$, $0 \leqq t \leqq s \leqq 1$, *of the integral equation*

$$e(t,s)+\int_0^s d(t,y,s)Q(s,y)\,dy+Q(s,t)=0.$$

Then $\lim_{n\to\infty} Q_n(t,s)=Q(t,s)$ *uniformly for* $0\leqq t\leqq s\leqq 1$.

Proof. We show first that if we define

$$\|Q_n-Q\|=\left[\int_0^s \{Q_n(s,t)-Q(s,t)\}^2\,dt\right]^{1/2},$$

then $\lim_{n\to\infty}\|Q_n-Q\|=0$ uniformly in s, $0\leqq s\leqq 1$. To do this we subtract the second integral equation above from the first to obtain the sequence of integral equations

$$\begin{aligned}0=e_n(t,s)-e(t,s)+\int_0^s d_n(t,y,s)[Q_n(s,y)-Q(s,y)]\,dy+Q_n(s,t)-Q(s,t)\\ +\int_0^s [d_n(t,y,s)-d(t,y,s)]Q(s,y)\,dy, \qquad n=1,2,\cdots.\end{aligned}$$

We assume that $\|Q_n-Q\|\leqq 1$ for $n=1,2,\cdots$. This is without loss, for if $\|Q_n-Q\|>1$ for some n we divide the above equation for that same n by $\|Q_n-Q\|$. This would change $e_n(t,s)-e(t,s)$ and $d_n(t,y,s)-d(t,y,s)$ but not the convergence properties of the sequences $\{e_n(t,s)-e(t,s)\}_{n=1}^{\infty}$ and $\{d_n(t,y,s)-d(t,y,s)\}_{n=1}^{\infty}$.

Let D_n^s be the linear operator

$$D_n^s f=\int_0^s d_n(t,y,s)f(y)\,dy, \qquad n=1,2,\cdots,$$

and D^s be the linear operator

$$D^s f=\int_0^s d(t,y,s)f(y)\,dy,$$

all defined for $f\in L^2(0,s)$. Since $\lim_{n\to\infty} e_n(t,s)=e(t,s)$ uniformly for $0\leqq t\leqq s\leqq 1$ and $\lim_{n\to\infty} d_n(t,y,s)=d(t,y,s)$ uniformly for $0\leqq t,\ y\leqq s\leqq 1$ we can conclude from the above sequence of integral equations that $\lim_{n\to\infty}[D_n^s(Q_n-Q)+(Q_n-Q)]=0$ uniformly in t and s for $0\leqq t\leqq s\leqq 1$. Further, it can be shown that $\{D_n^s(Q_n-Q)\}_{n=1}^{\infty}$ is a uniformly bounded, equi-continuous set of functions. Hence, there exists a subsequence $\{D_{n_i}^s(Q_n-Q)\}_{i=1}^{\infty}$ and a continuous function $\gamma(t,s)$, $0\leqq t\leqq s\leqq 1$, such that $\lim_{i\to\infty} D_{n_i}^s(Q_{n_i}-Q)=\gamma(t,s)$ uniformly for $0\leqq t\leqq s\leqq 1$. In addition the set of inequalities $\|D_{n_i}^s(Q_{n_i}-Q)+Q_{n_i}-Q\|+\|D_{n_i}^s(Q_{n_i}-Q)-\gamma\|\geqq\|\gamma-(Q_{n_i}-Q)\|$ implies that $\lim_{i\to\infty}\|\gamma-(Q_{n_i}-Q)\|=0$ uniformly for $0\leqq s\leqq 1$. And the set of inequalities

$$\begin{aligned}\|D^s\gamma+\gamma\|\leqq\|D\gamma-D_{n_i}\gamma\|+\|D_{n_i}\gamma-D_{n_i}(Q_{n_i}-Q)\|\\ +\|D_{n_i}(Q_{n_i}-Q)+(Q_{n_i}-Q)\|+\|(Q_{n_i}-Q)-\gamma\|\end{aligned}$$

implies that $\|D^s\gamma+\gamma\|=0$.

This last inequality says that $\gamma(s, t)=0$, $0\leqq t\leqq s\leqq 1$, since we have a unique solution for $D^s\gamma+\gamma=0$ by hypothesis. Hence $\lim_{i\to\infty}\|Q_{n_i}-Q\|=0$ uniformly in s, $0\leqq s\leqq 1$. Since Q is the unique solution to the integral equation $e(s,t)+\int_0^s d(t, y, s)Q(s, y)\,dy+Q(s,t)=0$ we have that $\lim_{n\to\infty}\|Q_n-Q\|=0$ uniformly for $0\leqq s\leqq 1$.

It remains to show that we have uniform pointwise convergence, that is, that $\lim_{n\to\infty} Q_n-Q=0$ uniformly for $0\leqq t\leqq s\leqq 1$. This follows easily from the fact that $\lim_{n\to\infty} D_n^s(Q_n-Q)+Q_n-Q=0$ uniformly for $0\leqq t\leqq s\leqq 1$ and from the fact that $\lim_{n\to\infty}\|Q_n-Q\|=0$ uniformly for $0\leqq s\leqq 1$.

THEOREM 5. *Let $g(s,t)$ and $h(s,t)$ have continuous partial derivatives up to order n (n included), ($n\geqq 0$), with respect to s and t for $0\leqq t\leqq s\leqq 1$. Suppose that the homogeneous equation*

$$\int_0^s h(t, y)k(y)+k(t)=0$$

has only the trivial solution $k(t)=0$, $0\leqq t\leqq s$. Then the unique solution $k(s,t)$ of the integral equation

$$g(s,t)+\int_0^s h(t,y)k(s,y)+k(s,t)=0$$

has continuous partial derivatives up to order n (n included) with respect to s and t.

Proof. From the Fredholm theory, see e.g. [16], we know that for fixed s, $k(s,t)$ has continuous partial derivatives up to order n (n included) derivatives in t, $0\leqq t\leqq s$. The remaining derivatives follow from the lemma on pp. 273–274 of [4].

We can now combine all of the above results to show that when the positive sequences $\lambda_1<\lambda_2<\cdots$ and $\rho_1,\rho_2,\cdots$ have the appropriate asymptotic forms, and if $f_n(s,t)$, $K_n(s,t)$, $n=1,2,\cdots$, are defined as in the Introduction, then

$$f(s,t)=\sum_{i=1}^{\infty}\left[\frac{Z_{\lambda_i}(s)Z_{\lambda_i}(t)}{\rho_i}-\frac{Z_{\lambda_i^*}(s)Z_{\lambda_i^*}(t)}{\rho_i^*}\right]$$

is a well-defined $C^4(0\leqq t\leqq s\leqq 1)$ function, and the integral equation

$$f(s,t)+\int_0^s f(t,y)K(s,y)\,dy+K(s,t)=0$$

has a unique solution $K(s,t)\in C^4(0\leqq t\leqq s\leqq 1)$ with the property that

$$\lim_{n\to\infty}\frac{\partial^{j+k}}{\partial s^j\,\partial t^k}K_n(s,t)=\frac{\partial^{j+k}}{\partial s^j\,\partial t^k}K(s,t)$$

for $j,k=0,1,2,3,4$, $0\leqq j+k\leqq 4$, uniformly for $0\leqq t\leqq s\leqq 1$. We prove the following theorem.

THEOREM 6. *Let $\lambda_1<\lambda_2<\cdots$ and ρ_1,ρ_2 be two positive sequences. Let Z_λ, λ_i^*, ρ_i^*, $i=1,2,\cdots$, and $f_n(s,t)$, $K_n(s,t)$, $n=1,2,\cdots$, be given as in the Introduction. Suppose that $(\lambda_i)^{1/4}=(\lambda_i^*)^{1/4}+P_i$ and $1/(\lambda_i\rho_i)=1/(\lambda_i^*\rho_i^*)+R_i$ and that $\sum_{i=1}^{\infty}(\lambda_i)^{k/4}|P_i|<\infty$ and $\sum_{i=1}^{\infty}(\lambda_i)^{k/4}|R_i|<\infty$ for $k=0,1,2,3,4$. Let $f(s,t)$ be given as in the Introduction. Then there exists a unique solution $K(s,t)\in$*

$C^4(0 \leqq t \leqq s \leqq 1)$ *of the integral equation*

$$f(s, t)+\int_0^s f(t, y)K(s, y)\,dy+K(s, t)=0$$

with the property that

$$\lim_{n\to\infty} \frac{\partial^{j+k}}{\partial s^j\,\partial t^k}K_n(s, t)=\frac{\partial^{j+k}}{\partial s^j\,\partial t^k}[K(s, t)]$$

uniformly for $0 \leqq t \leqq s \leqq 1$, $j, k = 0, 1, 2, 3, 4$, $0 \leqq j+k \leqq 4$.

Proof. We shall show first that there exists a unique solution $K(s, t)$ continuous in t, $0 \leqq t \leqq s$ for each s, $0 \leqq s \leqq 1$, of the equation

$$f(s, t)+\int_0^s f(t, y)K(s, y)\,dy+K(s, t)=0.$$

To this end, we recall that from Theorem 1 we have that the sum

$$\sum_{i=1}^{\infty}\left[\frac{Z_{\lambda_i}(s)Z_{\lambda_i}(t)}{\rho_i}-\frac{Z_{\lambda_i^*}(s)Z_{\lambda_i^*}(t)}{\rho_i^*}\right]$$

converges uniformly to the continuous function $f(s, t)$, $0 \leqq t \leqq s \leqq 1$. Further, from the remark following Corollary 2 of Theorem 2, we know that the set of functions $\{Z_{\lambda_i}(s)\}_{i=1}^{\infty}$ is complete on $L^2(0, 1)$. Hence Theorem 3 gives us the required result. In addition we can state that the derivatives $\partial^{j+k}K(s, t)/\partial s^j\,\partial t^k$ exist and are continuous for $j, k = 0, 1, 2, 3, 4$, $0 \leqq j+k \leqq 4$, and for $0 \leqq t \leqq s \leqq 1$. This follows directly from Theorems 1 and 5.

To show that the limits exist as in the statement of our theorem, we first recall that, as in the Introduction, $K_n(s, t)$ is the unique solution of

$$f_n(s, t)+\int_0^s f_n(t, y)K_n(s, y)\,dy+K_n(s, t)=0$$

where

$$f_n(s, t)=\sum_{i=1}^{n}\frac{Z_{\lambda_i}(s)Z_{\lambda_i}(t)}{\rho_i}-\frac{Z_{\lambda_i^*}(s)Z_{\lambda_i^*}(t)}{\rho_i^*}$$

and that $f_n(s, t)$, $K_n(s, t) \in C^{\infty}(0 \leqq t \leqq s \leqq 1)$. Further since $f_n(s, t) \to f(s, t)$ uniformly for $0 \leqq t \leqq s \leqq 1$, we have from Theorem 4 that $K_n(s, t) \to K(s, t)$ uniformly in t and s for $0 \leqq t \leqq s \leqq 1$. Also, once $K_n(s, t)$ is known, the derivative $(\partial/\partial s)K_n(s, t)$ may be determined as the unique solution, see [15], of the integral equation

$$\frac{\partial}{\partial s}f_n(s, t)+f_n(s, t)K_n(s, s)+\int_0^s f_n(t, y)\left[\frac{\partial}{\partial s}K_n(s, y)\right]dy+\frac{\partial}{\partial s}[K_n(s, t)]=0.$$

Since $(\partial/\partial s)f(s, t)$ and $(\partial/\partial s)K(s, t)$ exist and are continuous for $0 \leqq t \leqq s \leqq 1$ we may determine the derivative $(\partial/\partial s)K(s, t)$ as the unique solution of the integral equation

$$\frac{\partial}{\partial s}f(s, t)+f(s, t)K(s, s)+\int_0^s f(t, y)\left[\frac{\partial}{\partial s}K(s, y)\right]dy+\frac{\partial}{\partial s}[K(s, t)]=0.$$

From Theorem 1 we have that $\lim_{n\to\infty} (\partial/\partial s)f_n(s, t) = (\partial/\partial s)f(s, t)$ uniformly, and that $f_n(s, t) \to f(s, t)$ uniformly. We have previously applied Theorem 4 to obtain that $K_n(s, t) \to K(s, t)$ uniformly. Hence we apply Theorem 4 again to achieve that $\lim_{n\to\infty} (\partial/\partial s)K_n(s, t) = (\partial/\partial s)K(s, t)$ uniformly for $0 \leqq t \leqq s \leqq 1$. Similar proofs enable us to show that $\lim_{n\to\infty} (\partial^j/\partial s)K_n(s, t) = (\partial^j/\partial s^j)K(s, t)$, $j = 2, 3, 4$ uniformly for $0 \leqq t \leqq s \leqq 1$.

The remaining uniform limits, i.e.

$$\lim_{n\to\infty} \frac{\partial^{j+k}}{\partial s^j\, \partial t^k} K_n(s, t) = \frac{\partial^{j+k}}{\partial s^j\, \partial t^k} K(s, t),$$

uniformly for $0 \leqq t \leqq s \leqq 1$, $j = 0, 1, 2, 3, 4$, $k = 1, 2, 3, 4$, $0 \leqq j + k \leqq 4$, are shown more directly. Consider, first, the first partial derivative with respect to t. As in the above argument we can differentiate the integral equation satisfied by $K_n(s, t)$ to obtain that

$$\frac{\partial}{\partial t} K_n(s, t) = -\frac{\partial}{\partial t} f_n(s, t) - \int_0^s \left[\frac{\partial}{\partial t} f_n(t, y)\right] K_n(s, y)\, dy.$$

In addition we differentiate the integral equation satisfied by $K(s, t)$, with respect to t, to obtain

$$\frac{\partial}{\partial t} K(s, t) = -\frac{\partial}{\partial t} f(s, t) - \int_0^s \left[\frac{\partial}{\partial t} f(t, y)\right] K(s, y)\, dy.$$

Since $K_n(s, t) \to K(s, t)$ uniformly, as was proved above, and from Theorem 1 we have $(\partial/\partial t)f_n(t, y) \to (\partial/\partial t)f(t, y)$ uniformly and $(\partial/\partial t)f_n(s, t) \to (\partial/\partial t)f(s, t)$ uniformly, we must have $(\partial/\partial t)K_n(s, t) \to (\partial/\partial t)K(s, t)$ uniformly for $0 \leqq t \leqq s \leqq 1$. Similar proofs apply to the remaining derivatives.

Having shown that $K_n(s, t)$ and its first four derivatives converge uniformly to $K(s, t)$ and its corresponding derivatives, we are now prepared to prove our main result. That is, suppose we have the two positive sequences $\lambda_1 < \lambda_2 < \cdots$ and $\rho_1, \rho_2, \cdots$ satisfying the appropriate asymptotic forms. And suppose that $A_n(s)$ and $B_n(s)$ are coefficients, found as in [15], with the following properties: the real numbers $\lambda_1, \lambda_2, \cdots, \lambda_n, \lambda^*_{n+1}, \lambda^*_{n+2}, \cdots$ are all the eigenvalues for the eigenvalue problem

$$y^{(4)} + (A_n y^{(1)})^{(1)} + B_n y - \lambda y = 0,$$
$$y(0) = y^{(1)}(0) = y(1) = y^{(1)}(1) = 0;$$

if y_i^n is the eigenfunction corresponding to the ith eigenvalue, and if $(d^2/ds^2)[y_i^n]|_{s=0} = 1$, then $\int_0^1 [y_i^n(s)]^2\, ds = \rho_i$, $i = 1, 2, \cdots, n$ or $\int_0^1 [y_i^n(s)]^2\, ds = \rho_i^*$, $i = n+1, n+2, \cdots$; and if $K_n(s, t)$ is the unique $C^\infty(0 \leqq t \leqq s \leqq 1)$ solution of $f_n(s, t) + \int_0^s f_n(t, y) K_n(s, y)\, dy + K_n(s, t) = 0$, then we define the coefficients $A_n(s)$ and $B_n(s)$ as

$$A_n(s) = -4 \frac{d}{ds} K_n(s, s),$$

$$B_n(s) = -A_n(s) K_{n,s}(s, t)|_{t=s} + 2(K_{n,ss} - K_{n,tt})_t|_{t=s} - 2 \frac{d^3}{ds^3} K_n(s, s),$$

where $K_n(s,t)$ and its derivatives can be found by solving appropriate linear nonhomogeneous systems of equations. Then it will be shown in Theorem 7 that there exist functions $A(s)$ and $B(s)$ and $y_i(s)$, $i=1,2,\cdots$, such that $\lim_{n\to\infty} A_n(s)=A(s)$, $\lim_{n\to\infty} B_n(s)=B(s)$, $\lim_{n\to\infty} y_i^n(s)=y_i(s)$, $i=1,2,\cdots$, and such that $y_i(s)$ satisfies the differential equation

$$y_i^{(4)}+(Ay_i^{(1)})^{(1)}+By_i-\lambda_i y_i=0,$$

$i=1,2,\cdots$. To complete the analysis we need to show that $y_i(0)=y_i^{(1)}(0)=y_i(1)=y_i^{(1)}(1)=0$, $y_i^{(2)}(1)=1$, that $\lambda_1<\lambda_2<\cdots$ are *all* the eigenvalues for the eigenvalue problem

$$y^{(4)}+(Ay^{(1)})^{(1)}+By-\lambda y=0,$$
$$y(0)=y^{(1)}(0)=y(1)=y^{(1)}(1)=0$$

and that $\rho_i=\int_0^1[y_i(s)]^2\,ds$. This last analysis is done in § 3.

We proceed with Theorem 7.

THEOREM 7. *Let* Z_λ, λ_i^*, ρ_i^*, $K_n(s,t)$, $y_i^n(s)$, $i,n=1,2,\cdots$, *be given as in the Introduction. Suppose that the positive sequences* $\lambda_1<\lambda_2<\cdots$ *and* $\rho_1,\rho_2,\cdots$ *satisfy the conditions* $(\lambda_i)^{1/4}=(\lambda_i^*)^{1/4}+P_i$ *and* $1/(\lambda_i\rho_i)=1/(\lambda_i^*\rho_i^*)+R_i$ *where* $\sum_{i=1}^\infty(\lambda_i)^{k/4}|P_i|<\infty$ *and* $\sum_{i=1}^\infty(\lambda_i)^{k/4}|R_i|<\infty$, $k=0,1,2,3,4$. *Let* $K(s,t)\in C^4(0\leqq t\leqq s\leqq 1)$ *be defined as in Theorem* 6 *and let*

$$y_i(s)=Z_{\lambda_i}(s)+\int_0^s K(s,t)Z_{\lambda_i}(t)\,dt,$$

$$y_i^n(s)=\begin{cases} Z_{\lambda_i}(s)+\int_0^s K_n(s,t)Z_{\lambda_i}(t)\,dt, & i=1,2,\cdots,n,\\ Z_{\lambda_i^*}(s)+\int_0^s K_n(s,t)Z_{\lambda_i^*}(t)\,dt, & i=n+1,n+2,\cdots.\end{cases}$$

Let $A_n(s)$, $B_n(s)\in C^\infty[0,1]$ *be defined as above and let*

$$A(s)=-4\frac{d}{ds}K(s,s),$$

$$B(s)=-A(s)K_s(s,t)|_{t=s}+2(K_{ss}-K_{tt})_t|_{t=s}-2\frac{d^3}{ds^3}K(s,s).$$

Then $A(s)\in C^3[0,1]$, $B(s)\in C^1[0,1]$, $\lim_{n\to\infty}A_n(s)=A(s)$ *uniformly for* $0\leqq s\leqq 1$, $\lim_{n\to\infty}B_n(s)=B(s)$ *uniformly for* $0\leqq s\leqq 1$, $y_i(s)\in C^4[0,1]$,

$$\lim_{n\to\infty}\frac{d^k[y_i^n(s)]}{ds^k}=\frac{d^k}{ds^k}[y_i(s)],$$

uniformly for $0\leqq s\leqq 1$, $k=0,1,2,3,4$, *and*

$$y_i^{(4)}+(Ay_i^{(1)})^{(1)}+By_i-\lambda_i y_i=0,\qquad i=1,2,\cdots.$$

Proof. From Theorem 6 we have that all of the derivatives, up to and including the order four, of $K(s, t)$ exist and are continuous on $0 \leqq t \leqq s \leqq 1$. Hence $A(s)$ and $B(s)$ are well defined with $A(s) \in C^3[0, 1]$ and $B(s) \in C^1[0, 1]$. Further

$$\lim_{n\to\infty} \frac{\partial^{j+k}}{\partial s^j \, \partial t^k} K_n(s, t) = \frac{\partial^{j+k}}{\partial s^j \, \partial t^k} K(s, t)$$

uniformly for $0 \leqq t \leqq s \leqq 1$. Hence $\lim_{n\to\infty} A_n(s) = A(s)$, $\lim_{n\to\infty} (d/ds)A_n(s) = (d/ds)A(s)$, $\lim_{n\to\infty} B_n(s) = B(s)$ uniformly for $0 \leqq s \leqq 1$.

It remains to show the desired conclusions about $y_i(s)$, $i = 1, 2, \cdots$. First of all, since $K(s, t) \in C^4(0 \leqq t1)$, we have that $y_i(s) \in C^4[0, 1]$. To show the remaining conclusions, we shall make a slight change in notation and write

$$y_i^n = Z_{\lambda i} + \int_0^s K_n(s, t) Z_{\lambda_i}(t) \, dt, \qquad i = 1, 2, \cdots.$$

This is without loss since it is a true statement for $i \leqq n$ and we shall be concerned with limiting values as $n \to \infty$ for fixed i. Once this is done, we use the fact, from Theorem 6, that

$$\lim_{n\to\infty} \frac{\partial^{j+k}}{\partial s^j \, \partial t^k} K_n(s, t) = \frac{\partial^{j+k}}{\partial s^j \, \partial t^k} K(s, t)$$

uniformly for $0 \leqq t \leqq s \leqq 1$, $j, k = 0, 1, 2, 3, 4$, $0 \leqq j + k \leqq 4$, to show that $\lim_{n\to\infty} ((d^k/ds^k)[y_i^n(s)]) = (d^k/ds^k)[y_i(s)]$ uniformly for $0 \leqq s \leqq 1$, $k = 0, 1, 2, 3, 4$. Having this fact we may write

$$\begin{aligned} 0 &= \lim_{n\to\infty} \{(y_i^n)^{(4)} + [A_n (y_i^n)^{(1)}]^{(1)} + B_n y_i^n - \lambda_i y_i^n\} \\ &= (y_i)^{(4)} + (A y_i^{(1)})^{(1)} + B y_i - \lambda_i y_i, \end{aligned}$$

the last conclusion of the theorem.

3. Showing $\{y_i\}_{i=1}^{\infty}$ is a complete set of eigenfunctions. We have shown in § 2 that if we have two positive sequences, $\lambda_1 < \lambda_2 < \cdots$, $\rho_1, \rho_2, \cdots$ with given asymptotic forms, then we can find coefficients $A(s) \in C^3[0, 1]$, $B(s) \in C^1[0, 1]$ and functions $y_i(s) \in C^4[0, 1]$, $i = 1, 2, \cdots$, such that

$$y_i^{(4)} + (A y_i^{(1)})^{(1)} + B y_i - \lambda_i y_i = 0.$$

We seek now to show that $\lambda_1 < \lambda_2 < \cdots$ are all the eigenvalues for the eigenvalue problem

$$\begin{gathered} y^{(4)} + (A y^{(1)})^{(1)} + B y - \lambda y = 0, \\ y(0) = y^{(1)}(0) = y(1) = y^{(1)}(1) = 0. \end{gathered}$$

It will also be shown that y_i is the eigenfunction associated with λ_i, that $y_i^{(2)}(0) = 1$ and that $\rho_i = \int_0^1 [y_i(s)]^2 \, ds$. Our proof will consist of showing that $\{y_i\}_{i=1}^{\infty}$ is a complete, orthogonal set and that each y_i satisfies the five boundary conditions and has the right normalization constant.

THEOREM 8. *Let Z_λ, λ_i^*, ρ_i^*, $i = 1, 2, \cdots$, and $f_n(s, t)$, $K_n(s, t)$, $n = 1, 2, \cdots$, be given as in the Introduction. Suppose that the positive sequences $\lambda_1 < \lambda_2 < \cdots$ and*

$\rho_1, \rho_2, \cdots$ *satisfy the conditions* $(\lambda_i)^{1/4} = (\lambda_i^*)^{1/4} + P_i$ *and* $1/(\lambda_i \rho_i) = 1/(\lambda_i^* \rho_i^*) + R_i$ *where* $\sum_{i=1}^{\infty} (\lambda_i)^{k/4} |P_i| < \infty$ *and* $\sum_{i=1}^{\infty} (\lambda_i)^{k/4} |R_i| < \infty$, $k = 0, 1, 2$. *Let* $K(s, t)$ *be the unique continuous solution of*

$$f(s, t) + \int_0^s f(t, y) K(s, y)\, dy + K(s, t) = 0.$$

Let $y_i(s) = Z_{\lambda_i}(s) + \int_0^s K(s, t) Z_{\lambda_i}(t)\, dt$, $i = 1, 2, \cdots$. *Then* $\{y_i\}_{i=1}^{\infty}$ *is a complete, orthogonal sequence with* $\rho_i = \int_0^1 [y_i(s)]^2\, ds$ *and* $y_i(0) = y_i^{(1)}(0) = y_i(1) = y_i^{(1)}(1) = 0$, $y_i^{(2)}(0) = 1$, $i = 1, 2, \cdots$.

Proof. We discuss first the proof of the fact that $\{y_i\}_{i=1}^{\infty}$ is a complete, orthogonal set. The proof is the same as that in the first part of Theorem 1 of [15] provided that we can interchange integrations (with respect to s or t) with the infinite sum which defines $f(s, t)$. All that is needed then is that the infinite series

$$f(s, t) = \sum_{i=1}^{\infty} \left[\frac{Z_{\lambda_i}(s) Z_{\lambda_i}(t)}{\rho_i} - \frac{Z_{\lambda_i^*}(s) Z_{\lambda_i^*}(t)}{\rho_i^*} \right]$$

converges uniformly for $0 \leqq t \leqq s \leqq 1$. It is seen from the proof of Theorem 1 of this paper that the conditions on λ_i, ρ_i, $i = 1, 2, \cdots$, given above insure the required uniform convergence.

We next show that $\rho_i = \int_0^1 [y_i(s)]^2\, ds$. This follows directly from Theorem 7. For if $\lim_{n\to\infty} y_i^n(s) = y_i(s)$ uniformly for $0 \leqq s \leqq 1$ as was proved in Theorem 7 and if $\int_0^1 [y_i^n(s)]^2\, ds = \rho_i$ for $n \geqq i$ as is true by hypothesis, then

$$\rho_i = \lim_{n\to\infty} \rho_i = \lim_{n\to\infty} \int_0^1 [y_i^n(s)]^2\, ds = \int_0^1 \lim_{n\to\infty} [y_i^n(s)]^2\, ds = \int_0^1 [y_i(s)]^2\, ds.$$

It remains to show that the desired boundary conditions are achieved. The hypothesized asymptotic forms for λ_i and ρ_i insure that the first and second derivatives of $K(s, t)$ exist and are continuous. Hence the left boundary conditions can be verified directly recalling that $Z_\lambda(0) = Z_\lambda^{(1)}(0) = 0$ and $Z_\lambda^{(2)}(0) = 1$. The right boundary conditions are verified by recalling that $\lim_{n\to\infty} y_i^n(s) = y_i(s)$, $0 \leqq s \leqq 1$, and observing that also $\lim_{n\to\infty} (d/ds)[y_i^n(s)] = (d/ds)[y_i(s)]$, $0 \leqq s \leqq 1$.

Remark. The asymptotic forms required in Theorem 8 are somewhat weaker than those required in Theorem 1 and Theorem 7. The reason for this is that in order to prove the results of Theorem 8, we do not need *all* of the first four derivatives of $K_n(s, t)$ converging uniformly to the corresponding derivatives of $K(s, t)$. Accordingly, to show that $\{y_i\}_{i=1}^{\infty}$ is a complete, orthogonal set, that $\rho_i = \int_0^1 [y_i(s)]^2\, ds$, and that $y_i(0) = y_i(1) = 0$, $i = 1, 2, \cdots$, we need only require that $\sum_{i=1}^{\infty} |P_i| < \infty$. To show that $y_i^{(1)}(0) = y_i^{(1)}(1) = 0$, $i = 1, 2, \cdots$, we need that $\sum_{i=1}^{\infty} (\lambda_i)^{1/4} |P_i| < \infty$ and that $\sum_{i=1}^{\infty} (\lambda_i)^{1/4} |R_i| < \infty$. Finally to show that $y^{(2)}(0) = 1$ we use the hypothesis that $\sum_{i=1}^{\infty} (\lambda_i)^{1/2} |P_i| < \infty$, $\sum_{i=1}^{\infty} (\lambda_i)^{1/2} |R_i| < \infty$.

4. Other boundary conditions. We have proved our results requiring that Z_λ be defined for $\lambda > 0$ by

$$Z_\lambda^{(4)} - \lambda Z_\lambda = 0,$$

$$Z_\lambda(0) = Z_\lambda^{(1)}(0) = Z_\lambda(1) = 0, \qquad Z_\lambda^{(2)}(0) = 1.$$

Further we have let λ_i^*, ρ_i^*, $i=1,2,\cdots$, be the eigenvalues and normalization constants for the eigenvalue problem

$$Z^{(4)}-\lambda Z=0,$$
$$Z(0)=Z^{(1)}(0)=Z(1)=Z^{(1)}(1)=0, \qquad Z^{(2)}(0)=1;$$

and we have given our asymptotic conditions on $\lambda_1, \lambda_2, \cdots, \rho_1, \rho_2, \cdots$ in terms of the eigenvalues $\lambda_1^*, \lambda_2^*, \cdots$ and normalization constants $\rho_1^*, \rho_2^*, \cdots$.

We choose this particular problem to make the analysis easier but the analysis can be carried out under much more general circumstances. As is stated in Remark 2 at the end of § 2 of [15], the Z_λ's could be replaced by Y_λ's where for $\lambda>0$ each Y_λ satisfies the equation $Y_\lambda^{(4)}-\lambda Y_\lambda=0$ plus two homogeneous boundary conditions and one nonhomogeneous boundary condition at $s=0$ and one homogeneous boundary condition at $s=1$. That is, there would exist real $\alpha_i, \beta_i, \gamma_i, \eta_i$, $i=1,2,3,4$, independent of λ and with $\sum_{i=1}^4 \alpha_i^2 \neq 0$, $\sum_{i=1}^4 \beta_i^2 \neq 0$, $\sum_{i=1}^4 \gamma_i^2 \neq 0$, $\sum_{i=1}^4 \eta_i^2 \neq 0$ such that $\sum_{i=1}^4 \alpha_i Y_\lambda^{(i-1)}(0)=0$, $\sum_{i=1}^4 \beta_i Y_\lambda^{(i-1)}(0)=0$, $\sum_{i=1}^4 \gamma_i Y_\lambda^{(i-1)}(1)=0$, $\sum_{i=1}^4 \eta_i Y_\lambda^{(i-1)}(0)=1$. Furthermore, ρ_i^*, λ_i^*, $i=1,2,\cdots$, would be replaced by $\tilde{\rho}_i$, $\tilde{\lambda}_i$, $i=1,2,\cdots$, respectively, where $\tilde{\lambda}_i$, $i=1,2,\cdots$, is the entire set of eigenvalues, (and $\tilde{\rho}_i$ the corresponding set of normalization constants) for the eigenvalue problem

$$Y^{(4)}-\lambda Y=0,$$
$$0=\sum_{i=1}^4 \alpha_i Y^{(i-1)}(0)=\sum_{i=1}^4 \beta_i Y^{(i-1)}(0)=\sum_{i=1}^4 \gamma_i Y^{(i-1)}(1)=\sum_{i=1}^4 \delta_i Y^{i-1}(1),$$
$$\sum_{i=1}^4 \eta_i Y^{(i-1)}(0)=1,$$

(here δ_i is real, independent of λ, $i=1,2,3,4$, and $\sum_{i=1}^4 \delta_i^2 \neq 0$). It is assumed that $\tilde{\lambda}_i$ is simple, that the above eigenvalue problem is self-adjoint, and that for all μ, $\lambda>0$, Y_λ and Y_μ satisfy $[Y_\mu Y_\lambda^{(3)}-Y_\lambda Y_\mu^{(3)}-Y_\mu^{(1)}Y_\lambda^{(2)}+Y_\mu^{(2)}Y_\lambda^{(1)}]|_{s=0}=0$.

With all of these assumptions the analysis in [15] can be carried out to find coefficients and boundary conditions, so that $\lambda_1, \lambda_2, \cdots, \lambda_n, \tilde{\lambda}_{n+1}, \tilde{\lambda}_{n+2}, \cdots$ and $\rho_1, \rho_2, \cdots, \rho_n, \tilde{\rho}_{n+1}, \tilde{\rho}_{n+2}$ are all the eigenvalues and normalization constants for a fourth order, linear eigenvalue problem. We would then require that $\lambda_i^{1/4}=(\tilde{\lambda}_i)^{1/4}+P_i$ and $1/(\lambda_i\rho_i)=1/(\tilde{\lambda}_i\tilde{\rho}_i)+R_i$ with $\sum_{i=1}^\infty (\lambda_i)^{k/4}|P_i|<\infty$ and $\sum_{i=1}^\infty (\lambda_i)^{k/4}|R_i|<\infty$, $k=0,1,2,3,4$, to obtain the convergence of coefficients, and eigenfunctions as discussed in this paper. This leads to the existence of a fourth order eigenvalue problem whose entire set of eigenvalues is $\lambda_1, \lambda_2, \cdots$ and corresponding set of normalization constants is $\rho_1, \rho_2, \cdots$.

REFERENCES

[1] V. BARCILON, *On the solution of inverse eigenvalue problems of high orders*, Geophys. J. Roy. Astronom. Soc., 39 (1974), pp. 143–154.

[2] ———, *On the uniqueness of inverse eigenvalue problems*, Ibid., 38 (1974), pp. 287–298.

[3] G. BORG, *Eine Umkehrung der Sturm–Liouvilleschen Eigenvertaufgabe*, Acta Math., 78 (1946), pp. 1–96.

[4] I. M. GEL'FAND AND B. M. LEVITAN, *On the determination of a differential equation from its spectral function*, Izv. Akad. Nauk SSSR Ser. Mat., 15 (1951), pp. 309-360; Amer. Math. Soc. Transl., 1 (1955), pp. 253–304.
[5] K. KNOPP, *Theory of Functions, Part II*, Dover, New York, 1947.
[6] M. G. KREIN, *On a method of effective solution of an inverse boundary problem*, Dokl. Akad. Nauk SSSR, 94 (1954), pp. 987–990.
[7] ———, *Solution of the inverse Sturm–Liouville problem*, Ibid., 76 (1951), pp. 21–24.
[8] W. LEIGHTON AND Z. NEHARI, *On the oscillation of solutions of self-adjoint linear differential equations of the fourth order*, Trans. Amer. Math. Soc., 89 (1958), pp. 325–377.
[9] N. LEVISON, *The inverse Sturm–Liouville problem*, Mat. Tidsskr. B, (1949), pp. 25–30.
[10] ———, *Gap and Density Theorems*, Colloquium Publications, American Mathematical Society, New York, 1940.
[11] B. M. LEVITAN, *Generalized Translation Operators and Some of Their Applications*, Fizmatigz, Moscow, 1962; English trans., Israel Program for Scientific Translations, Jerusalem and Davey, New York, 1964.
[12] ———, *On the determination of a Sturm–Liouville Equation by two spectra*, Izv. Akad. Nauk SSSR Ser. Mat., 28 (1964), pp. 63–78; Amer. Math. Soc. Transl., 68 (1968), pp. 1–20.
[13] V. A. MARCENKO, *Concerning the theory of a differential operator of the second order*, Dokl. Akad. Nauk SSSR, 72 (1950), pp. 457–460.
[14] J. MCKENNA, *On the lateral vibration of conical bars*, SIAM J. Appl. Math., 21 (1971), pp. 265–278.
[15] J. MCLAUGHLIN, *An inverse eigenvalue problem of order four*, this Journal, 7 (1976), pp. 646–661.
[16] K. YOSIDA, *Lectures of Differential and Integral Equations*, Interscience, New York, 1960.

SIAM J. MATH. ANAL.
Vol. 9, No. 3, June 1978

Copyright © 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0903-0002$01.50/0

DILATIONS AS PROPAGATORS OF HILBERTIAN VARIETIES*

P. MASANI†

Abstract. In this paper W-to-W^* operator-valued positive definite kernels $K(\cdot\,\cdot)$ are defined on $\Lambda\times\Lambda$, where W is a Banach space and Λ an arbitrary set, and Hibertian varieties $X(\cdot)$ with such covariance kernels are studied with the aid of the Kolmogorov–Aronszajn–Pedrick kernel theorem. The general notion of a propagator of $X(\cdot)$ is introduced in terms of the action on Λ of a semi-group Γ, and necessary and sufficient conditions are established for its existence. We show that these conditions simplify substantially when the semi-group Γ is involutory, especially so when $\Lambda=\Gamma$. For $\Gamma=\Lambda$ equal to a unitized Banach algebra with isometric involution, these conditions subsume those given by Stinespring for C^*-algebras.

If for Hilbert spaces W and semi-groups Λ, dilations are redefined in terms of isometries rather than projections, then the dilation $\tilde{R}(\cdot)$ of a given W-to-W operator-valued function $R(\cdot)$ on Λ is precisely the propagator of a Hilbertian variety whose covariance kernel $K(\cdot\,\cdot)$ satisfies $K(\cdot,0)=R(\cdot)$. Dilation theorems are thus rendered explicit, and their method of proof routinized. From our results on propagators we deduce a simplified version of Nagy's principal theorem in which his translational inequality is mitigated, and the Bram version of Halmos's theorem on subnormal operators. Dilation theorems such as those of Lebow, Arveson and Naimark are shown to fit into this approach.

CONTENTS

1. Introduction. As usually enunciated, a simple *dilation theorem* asserts the embeddability of a Hilbert space W into a larger Hilbert space $\mathfrak{H}$ so that a given operator R on W to W can be retrieved by projection from a simpler operator $\tilde{R}$ on $\mathfrak{H}$ to $\mathfrak{H}$; more precisely,

$$RP=P\tilde{R}P, \quad \text{equivalently} \quad R=\mathrm{Rstr.}_W P\tilde{R}, \tag{1.1}$$

where P is the orthogonal projection on $\mathfrak{H}$ onto W, and "$\mathrm{Rstr.}_W$" stands for "restriction to W". Less simple versions assert the same type of embeddability for all positive powers R^n, $\tilde{R}^n$.

Dilation theorems are significant for physics and engineering in connection with systems whose inputs and outputs can be characterized by vectors on a Hilbert space W, and the action of the system by an operator R on W to W. Suppose, for instance, that this R is a contraction—the case of a dissipative filter. The theorem that contractions R have unitary dilations $\tilde{R}$ then shows that an equivalent filter can be obtained in the canonical way indicated in (1.1) from a nondissipative (ideal) one governed by $\tilde{R}$. This kind of knowledge contributes to the theory of systems designing, and may even suggest efficient and economic use of hardware.

An examination of the proofs of dilation theorems reveals however that in nearly every case, $\mathfrak{H}$ does not contain W but rather a subspace $\tilde{W}$ isometrically isomorphic to W; briefly, $W\sim\tilde{W}\subseteq\mathfrak{H}$ and $W\not\subseteq\mathfrak{H}$. Only when W is "identified" with $\tilde{W}$, does W "become" a subspace of $\mathfrak{H}$. Now identifications are compelling and useful in many mathematical situations, but in the present instance to identify W and $\tilde{W}$ is unnecessary

* Received by the editors October 23, 1975, and in revised form November 15, 1976.

† Department of Mathematics, University of Pittsburgh, Pittsburgh, Pennsylvania 15260. This work was supported by the National Science Foundation U.S.A. under Grants GP43072 and MPS 74-07302 A01.

and even misleading. For it obliterates the connection between the subject and the theory of propagators of varieties in Hilbert spaces or in spaces of Hilbert-space-operators which stem from positive definite kernels, as shown by theorems of a type initiated by Kolmogorov. This obscures the nature of the subject and cloaks it with a deceptive individuality.

In this paper we shall adopt a definition of dilation that is suggested by the proofs themselves, reveal the nexus between dilation theory and the theory of Hilbertian varieties and their propagators, and thereby appreciably simplify and systematize the former. Among our new results we would call attention to the Main Thms. 4.10, 4.13, from which emerge readily a substantially improved version of Nagy's principal theorem [26, p. 21] and the theorem of Stinespring [32, Thm. 1], and from which the Bram theorem [6, Thm. 1] follows in a clearcut way (cf. Thms. 5.3, 4.14, 4.15 and 6.2). Our definition reads as follows:

1.2 Definition. (a) Let $W, \mathfrak{H}$ be Hilbert spaces over the field $\mathbb{F}$, R be any operator (not necessarily linear) from W to W and $\tilde{R}$ be any operator from $\mathfrak{H}$ to $\mathfrak{H}$. We say that *$\tilde{R}$ is a dilation of R* iff $\exists$ a linear isometry J on W into $\mathfrak{H}$ such that $R = J^*\tilde{R}J$.

(b) Let Λ be any set, W and $\mathfrak{H}$ be Hilbert spaces over $\mathbb{F}$, $R(\cdot)$ be a W-to-W operator-valued function on Λ, and $\tilde{R}(\cdot)$ be a $\mathfrak{H}$-to-$\mathfrak{H}$ operator-valued function on Λ. We say that *$\tilde{R}(\cdot)$ is a dilation of $R(\cdot)$* iff $\exists$ a linear isometry J on W into $\mathfrak{H}$ such that

$$\forall \lambda \in \Lambda, \qquad R(\lambda) = J^*\tilde{R}(\lambda)J.$$

In this approach the equation (1.1) gives way to that in Def. 1.2(a).[1] A *dilation theorem* is one that claims for a given type of W-to-W operator R or W-to-W operator-valued function $R(\cdot)$, the existence of a dilation $\tilde{R}$ or $\tilde{R}(\cdot)$ of a specific type, in the sense of Def. 1.2. When perceived in this way, dilation theory comes to depend on the generalizations of the following theorem of Kolmogorov on positive definite kernels [15, Lemma 2]:[2]

1.3 Kolmogorov's Theorem. *Let $k(\cdot\,\cdot)$ be a positive definite* (PD) *kernel on $\mathbb{N}_+ \times \mathbb{N}_+$ to $\mathbb{F}$, and let $\mathfrak{H}$ be any infinite dimensional Hilbert space over $\mathbb{F}$. Then $\exists$ a sequence $(x_n)_1^\infty$ in $\mathfrak{H}$ such that*

$$\forall m, n \in \mathbb{N}_+, \qquad (x_m, x_n) = k(m, n).$$

This is the first in a line of theorems which assert the realizability of scalar-valued or operator-valued PD kernels on spaces $\Lambda \times \Lambda$ as covariance kernels of varieties in Hilbert spaces or in spaces of Hilbert-space-operators.[3] A culminating step in this development was the Moore–Aronszajn reproducing kernel theorem [2]. Very recently Allen, Narcowich and Williams [1] have extended the Kolmogorov result to W-to-W operator-valued PD kernels on $\Lambda \times \Lambda$, where W is any separable Hilbert space and Λ a separable Hausdorff space, by a direct proof free of the reproducing kernel technique, and in this paper, Thm. 2.10, we state its further extension to W-to-W^* operator-valued kernels, where W is any Banach space and W^* is its adjoint, using certain ideas

[1] Our use of a nonprojection J in 1.2 seems to be in keeping with recent trends, cf. the enunciation of Lebow's theorem in [5, Thm. 1.0] in which J is an isometry, and especially Stinespring [32, Thm. 1] in which J is any continuous linear operator.

[2] In this paper $\mathbb{F}$ will refer to either the real number field $\mathbb{R}$ or the complex number field $\mathbb{C}$, and $\mathbb{N}$ to the set of all integers. $\mathbb{N}_+$, $\mathbb{R}_+$, and $\mathbb{N}_{0+}$, $\mathbb{R}_{0+}$ will denote the subsets of positive elements, and subsets of nonnegative elements of $\mathbb{N}$ and $\mathbb{R}$.

[3] For specific Hilbert spaces such $L_2(\Omega, \mathcal{B}, P)$ and specific types of PD kernels, such theorems have a history going back at least to Khinchine's work on stationary stochastic processes [14] (1934).

of Vakhania [33]. These ideas appear, however, in the 1957 technical report [30][4] by G. B. Pedrick, in which W is just any locally convex topological vector space.

Theorems of the Kolmogorov–Aronszajn–Pedrick type permit the translation of problems depending on positive definiteness into ones about Hilbertian varieties, and thereby open the way to using knowledge about the structure of the latter to solve the former problems.[5] But such theorems are also crucial for dilation theory in view of the fundamental fact, first perceived by Halmos [11] and Nagy [26], that *a dilation-possessing operator-valued function* $R(\cdot)$ *on* Λ *gives rise to a* PD *kernel* $K(\cdot\cdot)$ *which reflects the specific properties of* $R(\cdot)$, *and differs from the trivial kernel* $J(\cdot\cdot)$ *on* $\Lambda\times\Lambda$ *given by* $J(\lambda,\lambda')=R(\lambda')^*\cdot R(\lambda)$. For instance, for the semi-group $(C^n: n\in\mathbb{N}_{0+})$ of linear contractions on a Hilbert space W, Nagy defined the kernel $K(\cdot\cdot)$ on $\mathbb{N}\times\mathbb{N}$ by:

$$K(m,n)=C(m-n),\qquad C(n)\underset{d}{=}\begin{cases}C^n, & n\geqq 0,\\ (C^{-n})^*, & n<0.\end{cases}$$

The nonobvious fact that this $K(\cdot\cdot)$ is PD rests on the C^n being contractions; cf. [26, p. 30], [27, p. 29].

Now it is well known to workers in random processes and random fields that Hilbertian varieties associated with PD kernels by theorems of the Kolmogorov–Aronszajn type often carry with them in a natural way certain operator-valued functions such as unitary groups, isometric semi-groups, spectral measures, etc. These operator-valued functions act as *propagators* or *controllers* of the varieties; cf. Def. 3.2(b). This holds in particular for all the nontrivial kernels considered by dilation theorists, and the propagators turn out to be precisely the desired dilations. One can in fact lay down the following routine for obtaining all dilations:

1.4 PROCEDURE. *Given the* W-to-W *operator-valued function* $R(\cdot)$ *on* Λ, (i) *find the appropriate* (*nontrivial*) PD *kernel* $K(\cdot\cdot)$ *on* $\Lambda\times\Lambda$ (*if any*), (ii) *find the vector- or operator-valued Hilbertian variety* $X(\cdot)$ *having the covariance kernel* $K(\cdot\cdot)$, (iii) *determine the operator-valued controller or propagator* $\tilde{R}(\cdot)$ *of this variety. Then* $\tilde{R}(\cdot)$ *is the dilation of* $R(\cdot)$, *and the isometry* J *in* 1.2 *is the value of* $X(\cdot)$ *at some distinguished point on* Λ.[6]

The verification of the last assertion in 1.4 just involves writing the relevant definitions and combining them trivially to get the equations in Def. 1.2 (see proof of Thm. 5.1 below). Of the three steps in 1.4, it is (i) which involves the hardest work. In the classical proofs steps (ii) and (iii) are not sharply distinguished, and the variety is

[4] As this report has remained unpublished, results subsumed in it have been rediscovered in subsequent years. The case in which W is a Hilbert space was treated by MacNerney in his 1960 paper [18], the contents of which was announced in 1955 in the abstracts [17]. We find that our Thm. 2.10 may be retrieved from Pedrick's Thm. 5 [30, p. 42]. But this is stated for W^*-to-W operator-valued kernels, and in an idiom which emphasizes the reproducing rather than covariance property. So it is easier to prove Thm. 2.10 directly from Aronszajn's Thm. (see Appendix C).

[5] Although this method is still new, it has shed new light on several important problems depending on positive definiteness. Among the notable examples: 1) the F. and M. Riesz theorem on bounded complex-valued measures on the unit circle C of $\mathbb{C}$, [31, 4.3]; 2) the theorem of factorization of a matrix-valued function on C in the form $\Phi(\cdot)\Phi(\cdot)^*$ where Φ is in the Hardy class on the inner disk D_+ [35] and its many extensions culminating in [25, Thm. 3.8]; 3) the factorization of a matrix-valued Hardy class function Ψ on D_+ into optimal and residual (Beurling's outer and inner) factors [19]; 4) the Levy–Khinchine theorem for infinitely decomposable probability distributions over $\mathbb{R}^q$ [22].

[6] In many instances, the dilations are known in stochastic and related work, and a search of the literature will often give this free.

encountered from the start with the propagator inserted.[7] This mix-up is confusing, for step (ii) can be taken for any set Λ, whereas step (iii) requires Λ to have some algebraic structure. In pioneering days this initial merger of steps (ii) and (iii) was perhaps unavoidable, since the generalization (2.10 below) of Aronszajn's theorem was unknown and the notion of propagator was somewhat amorphous. But today a more perfect architecture is revealed by the realization of 1.4 and of the close alliance between dilation theory and the theories of Hilbertian varieties and of their controllers and propagators.

In much the same spirit, the concept of *extension* to a larger Hilbert space $\mathfrak{H}$, as opposed to dilation in $\mathfrak{H}$, cf. [26, p. 19], also needs widening. The definition which naturally uggests itself is as follows:

1.5 DEFINITION. Let Λ, W, $\mathfrak{H}$, $R(\cdot)$, $\tilde{R}(\cdot)$ be as in 1.2(b). We say that $\tilde{R}(\cdot)$ is an *image-extension* of $R(\cdot)$ iff $\exists$ a linear isometry J on W into $\mathfrak{H}$ such that

$$\forall\lambda\in\Lambda,\qquad JR(\lambda)J^{-1}\subseteq\tilde{R}(\lambda),$$

i.e. such that $\tilde{R}(\lambda)$ is an extension of the image $JR(\lambda)J^{-1}$, in $\mathrm{CL}(\mathcal{R}_J, \mathcal{R}_J)$, of $R(\lambda)$, $\mathcal{R}_J \underset{d}{=}$ range of J.

In § 2 we define Hilbertian varieties and formulate a very general version of the kernel theorem following the thought of Kolmogorov, Aronszajn and Pedrick (Thm. 2.10). In this Λ is an arbitrary set and W is only a Banach space. We consider the continuity questions which arise when Λ is equipped with a Hausdorff topology, and the measurability questions which come up when Λ is equipped with a σ-algebra (Cors. 2.13, 2.15). We also study the case where Λ is a vector space and the kernel $K(\cdot\,\cdot)$ is sesquilinear (Cor. 2.16). As most of the material in § 2 is ancillary to dilation theory in the narrow sense and is formulated with somewhat greater generality than needed for dilation purposes, the proofs of all theorems in § 2 are relegated to Appendices at the end of the paper.

In § 3 we first define a wide concept of propagator or controller $S(\cdot)$ of a Hilbertian variety $X(\cdot)$ on Λ: it is only required that $S(\cdot)$ be the representation of a not necessarily Abelian semi-group Γ that acts on Λ. We then establish the necessary and sufficient conditions that $X(\cdot)$ should possess such an $S(\cdot)$ (Thm. 3.4). We also consider the propagators of stationary varieties (Cor. 3.5), and the strong continuity of the propagator when the semi-group Γ is topological (Cor. 3.6).

In § 4 we turn to the important case in which Γ is an involutory semi-group in the sense of Nagy [26, § 6]. Our main objective here is to show that the necessary and sufficient conditions for the existence of a propagator are much milder than those required in Thm. 3.4 for noninvolutory Γ (Main Thm. 4.10). The conditions are that the covariance kernel of the variety possess mild transfer and translational properties, (4.5a–b). As a first step we show that the transfer requirement (4.5a) alone is necessary and sufficient to yield closed, single-valued (but possibly discontinuous) propagators on domains having a common everywhere dense linear manifold (Thm. 4.7). Then with some clues from Stinespring's proof we show that the mild translational requirement (4.5b) secures their continuity as well. Our main Thm. 4.10 can be recast in terms of the notion of PD function $R(\cdot)$ on Γ due to Nagy (Thm. 4.13). Next we take the Γ in 4.13 to be a Banach algebra $\mathbb{A}$ with a unit and an isometric involution $*$ (viewing it as a

[7] For instance, the Gelfand-Raikov Thm. [9] that if $\phi(\cdot)$ is a PD function on a topological group Λ, then $\phi(t)=(U(t)\alpha,\alpha)$ where $U(\cdot)$ is a unitary representation of Λ over a Hilbert space $\mathfrak{H}$ and $\alpha\in\mathfrak{H}$, combines two assertions: 1) there is a variety in $\mathfrak{H}$ having the covariance $\phi(\cdot)$, 2) this variety being stationary, its propagator is a unitary group.

multiplicative $*$ s.g.), and the PD function $R(\cdot)$ on $\mathbb{A}$ to be a linear operator on $\mathbb{A}$, and obtain as a corollary a generalization of Stinespring's theorem for C^*-algebras [32, Thm. 1], (Thms. 4.14, 4.15).

In § 5 we assume that the parameter space Λ is itself a semi-group and W is a Hilbert space, and prove a general dilation theorem for all W-to-W operator-valued PD kernels for which the associated Hilbertian varieties have propagators, (Thm. 5.1). This validates Procedure 1.4, and shows that the only nonroutine aspect in proving a dilation theorem is the discovery of an appropriate kernel $K(\cdot\cdot)$ on $\Lambda\times\Lambda$ and the demonstration of both its positive definiteness and its fulfillment of the propagator existence requirements given in §§ 3,4. As an immediate corollary of Thms. 5.1 and 4.13 we get an explicit and simplified form of Nagy's principal theorem [26, p. 21] in which his translational requirement (c) is substantially mitigated (Thm. 5.3).[8] Another such corollary is a dilation-version of our extension of Stinespring's theorem [32, Thm. 1], (Thm. 5.4). From Thm. 5.4, Lebow's pioneering theorem [16, p. 84] involving spectral sets, which subsumes classical dilation theorems, can be retrieved by Arveson's method of extensions [5, 0.1].

Next, we show (Thm. 5.12) that in the classical Naimark dilation theorem for a bounded W-to-W nonnegative hermitian operator-valued measure, the Hilbertian variety is a *quasi-isometric measure*, (5.7), and the dilation is its *spatial spectral measure*, Def. 5.10, useful concepts which emerged from considerations quite independent of dilation theory; cf. [21]. A natural candidate for the dilation space is the space of W-valued functions on Λ which are "L_2" with respect to the nonnegative hermitian operator-valued measure M, (Remark 5.13). We thus provide an explicit version of the Naimark theorem.

In § 6 we turn to the work of Halmos and Bram on normal extensions in the new setting in which projections give way to isometries. Just as Nagy showed in [26, § 10] that the Halmos theorem could be deduced from his principal theorem, we now show that the substantially improved version of the Halmos theorem due to Bram [6, Thm. 1] follows from our simplified version (Thm. 5.3) of Nagy's theorem.

Initial parts of this paper bear the impress of some valuable conversations with Dr. V. Mandrekar in the summer of 1975 on Aronszajn's theorem and its possible extensions. The writer is happy to acknowledge that his conception of the scalar kernel associated to an operator-valued PD kernel (cf. 2.6(γ)) is especially traceable to these discussions. This was before we became aware of Pedrick's unpublished report [30]. For this awareness thanks are due to Dr. H. Salehi. The paper was subsequently revised and enlarged. The revised portions bear the benefit of conversations with Drs. J. Conway, P. Jørgensen and especially W. Arveson.[9]

2. The covariance kernel of a Hilbertian variety. In this section we shall introduce the concept of nonnegative hermitianness for linear operators from a Banach space W to its adjoint W^* and the concept of a W-to-W^* operator-valued positive definite kernel. Our definitions are modifications of those due to Vakhania [33] and Chobanian

[8] The writer is grateful to Dr. Arveson for communicating his (unpublished) observation that Nagy's premiss (c) is redundant for *bounded* PD functions, and a sketch of his proof. This information kindled the writer's interest in involutory semi-groups and led him to Thms. 4.10 and 5.3. The relationship between Arveson's observation and our Thm. 5.3 is not as yet clear. Arveson's proof appeals to Stinespring's theorem on C^*-algebras, whereas ours does not, and we get the latter theorem as a corollary.

[9] These informal but useful discussions occurred at the Conference on Operator Theory held at the University of New Hampshire in Durham in the summer of 1976, and the writer is very grateful to the organizers, Drs. E. Nordgren and B. Moore III, for their invitation.

[7] who define these ideas for *semi*-linear operators on W to its dual W'.[10] Our goal is to give a very general definition of Hilbertian variety and its covariance kernel, which will yield the requisite propagator theory (§§ 3, 4), and when W is a Hilbert space provide all the varieties encountered in dilation theory (§ 5). For the purposes of the latter theory we shall give a general formulation of the Kolmogorov, Aronszajn, Pedrick theorems for a Banach space W (Thm. 2.10). An even more general formulation in which W is only a locally convex topological vector space is possible as Pedrick has shown [30]. It will be understood that:

(2.1)
- (a) Λ *is any set,*
- (b) W *is a Banach space over* $\mathbb{F}$ ($\mathbb{F}=\mathbb{R}$ *or* $\mathbb{C}$),
- (c) W^* *is the adjoint of* W,
- (d) $\mathrm{CL}(\mathscr{X}, \mathscr{Y})$ *is the space of all continuous linear operators on* $\mathscr{X}$ *to* $\mathscr{Y}$, *where* $\mathscr{X}$, $\mathscr{Y}$ *are normed vector spaces,*
- (e) $\forall A \in \mathrm{CL}(\mathscr{X}, \mathscr{Y})$, A^* *is the adjoint of* A.

2.2 *Remarks.* (a) The distinction between the adjoint W^* and dual W' of the Banach space W is important here:

$$W^* \underset{d}{=} \{\overline{f(\cdot)}: f(\cdot) \in W'\},$$

i.e. W^* is the set of all continuous *semi-linear* functionals on W to $\mathbb{F}$. It follows that if $A \in \mathrm{CL}(W, W^*)$, then

$$\forall w, w' \in W \quad \text{and} \quad \forall c, c' \in \mathbb{F}, \qquad [A(cw)](c'w') = \bar{c}'c \cdot [A(w)](w').$$

Of course $W^* = W'$ when $\mathbb{F}=\mathbb{R}$.

(b) For $A \in \mathrm{CL}(\mathscr{X}, \mathscr{Y})$ the *adjoint* A^* of A is defined to be the $\mathscr{Y}^*$-to-$\mathscr{X}^*$ operator such that

$$A^*(f) = f \circ A, \qquad f \in \mathscr{Y}^*.$$

Thus $A^* \in \mathrm{CL}(\mathscr{Y}^*, \mathscr{X}^*)$, and A^* differ in general from the *dual* A' of A, which is a $\mathscr{Y}'$-to-$\mathscr{X}'$ operator. We have $(cA)^* = \bar{c}A^*$, $c \in \mathbb{F}$. However, $|A^*| = |A'| = |A|$.

(c) If $\mathfrak{H}$ is a Hilbert space over $\mathbb{F}$, then $\mathfrak{H}^*$ is isometrically isomorphic to $\mathfrak{H}$. Now *in this paper we identify* $\mathfrak{H}^*$ *and* $\mathfrak{H}$. Consequently, the adjoint J^* of an operator $J \in \mathrm{CL}(W, \mathfrak{H})$ is a $\mathfrak{H}$-to-W^* operator. Moreover from the definition of the adjoint given in (b) and $\mathfrak{H} = \mathfrak{H}^*$, it follows readily that null space $J^* = (\text{Range } J)^\perp$ in $\mathfrak{H}$, and if $J_1, J_2 \in \mathrm{CL}(W, \mathfrak{H})$ then

$$\forall w_1, w_2 \in W, \qquad [(J_2^* J_1)(w_1)](w_2) = (J_1 w_1, J_2 w_2)_{\mathfrak{H}}.$$

The term on the left supplants the familiar $(J_2^* J_1 w_1, w_2)_W$ for a Hilbert space W. It is also easy to see that

$$\forall J \in \mathrm{CL}(W, \mathfrak{H}), \qquad |J^*J| = |J|^2.$$

2.3 Definition. (a) An operator $H \in \mathrm{CL}(W, W^*)$ is called *hermitian* iff $\forall w, w' \in W$, $\{H(w')\}(w) = \overline{\{H(w)\}(w')}$.

(b) An operator $H \in \mathrm{CL}(W, W^*)$ is called *nonnegative* (in symbols $H \succeq 0$) iff H is hermitian and $\forall w \in W$, $\{H(w)\}(w) \geqq 0$.

(c) For $H_1, H_2 \in \mathrm{CL}(W, W^*)$ we write $H_1 \succeq H_2$ to mean that $H_1 - H_2 \succeq 0$.

Note. When W is a Hilbert space, the condition in 2.3(a) becomes: $\forall w, w' \in W$, $(Hw', w) = \overline{(Hw, w')}$. Since the last term is (w', Hw), this condition is just the usual one for hermitianness. Unlike the case of a Hilbert space W, a W-to-W^* hermitian operator

[10] English translations of several papers by these authors on this subject and its stochastic ramifications are available in a Michigan State University Technical Report by Dr. A. G. Miamee.

H is *not* self-adjoint, i.e. $H \neq H^*$. This is because H^* is on W^{**} to W^* and not on W to W^*. If W is reflexive and we identify W^{**} with W, then of course H becomes self-adjoint. As for the condition 2.3(b), it too reduces to the usual one for nonnegativity when W is a Hilbert space.

The proof of the following triviality is left to the reader. The parts (c), (d), (e), which are generalizations of the parallelogram law, the polarization identity and the generalized Schwarz inequality, are immediate consequences of the sesquilinearity of the kernel $\phi(\cdot\cdot)$ on $W \times W$ to $\mathbb{F}$, defined by $\phi(w, w') = R(w)(w')$, where $R \in \mathrm{CL}(W, W^*)$. Part (f) follows from (e), and (g) follows from (f).

2.4 TRIVIALITY. *Let $H, H_1, H_2 \in \mathrm{CL}(W, W^*)$ and $w, w' \in W$. Then*

(a) H_1, H_2 *are hermitian and* $c_1, c_2 \in \mathbb{R} \Rightarrow c_1H_1 + c_2H_2$ *is hermitian*;

(b) $H_1, H_2 \geq 0$ *and* $C_1, C_2 \in \mathrm{CL}(W, W) \Rightarrow C_1^*H_1C_1 + C_2^*H_2C_2 \geq 0$;

(c) $H(w+w')(w+w') + H(w-w')(w-w') = 2\{H(w)(w) + H(w')(w')\}$;

(d) $4H(w)(w') = \{H(w+w')(w+w') - H(w-w')(w-w')\}$
$+ i\{H(w+iw')(w+iw') - H(w-iw')(w-iw')\}$;

(e) $H \geq 0 \Rightarrow |H(w)(w')| \leqq \sqrt{\{H(w)(w)\}} \cdot \sqrt{\{H(w')(w')\}}$;

(f) $$H \geq 0 \Rightarrow |H| = \sup_{0 \neq w \in W} \frac{H(w)(w)}{|w|^2};$$

(g) $H_1 \geq H_2 \geq 0 \Rightarrow |H_1| \geqq |H_2|$.

From the Remarks 2.2 it is clear that the following extended definition of positive definite kernel is consistent with the standard one for scalar-valued kernels:

2.5 DEFINITION. A kernel $K(\cdot\cdot)$ on $\Lambda \times \Lambda$ to $\mathrm{CL}(W, W^*)$ is called *positive definite* (PD) iff $\forall$ functions $C(\cdot)$ on Λ to $\mathrm{CL}(W, W)$, $\forall r \in \mathbb{N}_+$, and $\forall \lambda_1, \cdots, \lambda_r \in \Lambda$,

$$\sum_{i=1}^{r} \sum_{i=1}^{r} C(\lambda_j)^* K(\lambda_i, \lambda_j) C(\lambda_i) \geq 0,$$

and[11]

$$\forall \lambda, \lambda' \in \Lambda \quad \text{and} \quad \forall w, w' \in W, \qquad [K(\lambda, \lambda')(w)](w') = \overline{[K(\lambda', \lambda)(w')](w)}.$$

The following lemma, giving conditions equivalent to positive definiteness, is required for the very formulation of our extension of the Moore–Aronszajn theorem.

2.6 MAIN LEMMA. *Let $K(\cdot\cdot)$ be a kernel on $\Lambda \times \Lambda$ to* $\mathrm{CL}(W, W^*)$. *Then the following conditions are equivalent*:

(α) $K(\cdot\cdot)$ *is PD*,

(β) $\forall$ *functions* $w(\cdot)$ *on* Λ *to* W, $\forall r \in \mathbb{N}_+$, *and* $\forall \lambda_1, \cdots, \lambda_r \in \Lambda$

$$\sum_{i=1}^{r} \sum_{j=1}^{r} [K(\lambda_1, \lambda_j)(w(\lambda_i))](w(\lambda_j)) \geqq 0$$

and[12]

$$\forall \lambda, \lambda' \in \Lambda \quad \textit{and} \quad \forall w, w' \in W, \qquad [K(\lambda, \lambda')(w)](w') = \overline{[K(\lambda', \lambda)(w')](w)},$$

(γ) *if* $\forall \lambda, \lambda' \in \Lambda$ *and* $\forall w, w' \in W$,

$$k\{\lambda, w), (\lambda', w')\} \underset{d}{=} [K(\lambda, \lambda')(w)](w'),$$

$k(\cdot\cdot)$ *is a* PD *kernel on* $(\Lambda \times W) \times (\Lambda \times W)$ *to* $\mathbb{F}$.

Proof. See Appendix A.

[11] When $\mathbb{F} = \mathbb{C}$ the condition to come follows from the last, and is therefore redundant. We leave this to the reader to check. When W is a Hilbert space, this conditon has the equivalent simpler rendering: $K(\lambda, \lambda') = K(\lambda', \lambda)^*$.

[12] See Footnote 11.

The following are some simple properties of operator-valued PD kernels, easily obtainable via 2.6(γ) from the familiar properties of $\mathbb{F}$-valued PD kernels:

2.7 TRIVIALITY. *Let* $K(\cdot\,\cdot)$ *be a* PD *kernel on* $\Lambda\times\Lambda$ *to* $\mathrm{CL}(W, W^*)$, *and* $\lambda, \lambda'\in\Lambda$ *and* $w, w'\in W$. *Then*

(a) $K(\lambda,\lambda)$ *is hermitian and nonnegative* (Def. 2.3), *and*

$$0\leqq[K(\lambda,\lambda)(w)](w)\leqq|K(\lambda,\lambda)|\,|w|^2,$$

(b) $|[K(\lambda,\lambda')(w)](w')\leqq\sqrt{[K(\lambda,\lambda)(w)](w)}\cdot\sqrt{[K(\lambda',\lambda')(w')](w')},$

(c) $|K(\lambda,\lambda')|\leqq\sqrt{|K(\lambda,\lambda)|}\cdot\sqrt{|K(\lambda',\lambda')|}.$

In a nutshell, our generalization of the Moore–Aronszajn theorem asserts that *the operator-valued* PD *kernel* $K(\cdot\,\cdot)$ *factors with variables separated via the reproducing kernel Hilbert space of the associated* $\mathbb{F}$*-valued* PD *kernel* $k(\cdot\,\cdot)$ *of* 2.6(γ). Before giving the full enunciation, it is necessary to define Hilbertian varieties, and dispose of the question of uniqueness up to unitary equivalence:

2.8 DEFINITION. (a) By a *Hilbertian variety* we shall mean a function $x(\cdot)$ on Λ to $\mathfrak{H}$ or a function $X(\cdot)$ on Λ to $\mathrm{CL}(W,\mathfrak{H})$ where $\mathfrak{H}$ is any Hilbert space.

(b) The *subspace* of a Hilbertian variety $x(\cdot)$ or $X(\cdot)$ is defined by[13]

$$\mathcal{S}_x \underset{d}{=} \mathfrak{S}\{x(\lambda)\colon \lambda\in\Lambda\},\qquad \mathcal{S}_X \underset{d}{=} \mathfrak{S}\{X(\lambda)(W)\colon \lambda\in\Lambda\}.$$

(c) By the *covariance kernel* of a Hilbertian variety $x(\cdot)$ or $X(\cdot)$ we shall mean the function $k(\cdot\,\cdot)$ or $K(\cdot\,\cdot)$ on $\Lambda\times\Lambda$ defined by

$$k(\lambda,\lambda')\underset{d}{=}(x(\lambda),x(\lambda'))_{\mathfrak{H}},\qquad K(\lambda,\lambda')\underset{d}{=}X(\lambda')^*\cdot X(\lambda).$$

Thus $k(\cdot\,\cdot)$ takes values in $\mathbb{F}$ and $K(\cdot\,\cdot)$ takes values in $\mathrm{CL}(W, W^*)$.

We now assert that any two Hilbertian varieties having the same covariance kernel are unitarily equivalent:

2.9 CONGRUENCE THEOREM. *Let* (i) W *be a Banach space and* $\mathfrak{H}$, $\mathfrak{K}$ *be Hilbert spaces over* $\mathbb{F}$, (ii) $X(\cdot)$, $Y(\cdot)$ *be functions on* Λ *to* $\mathrm{CL}(W,\mathfrak{H})$, $\mathrm{CL}(W,\mathfrak{K})$, *respectively, having the same covariance kernel, i.e. such that*

$$\forall\lambda',\lambda\in\Lambda,\qquad X(\lambda')^*X(\lambda)=Y(\lambda')^*Y(\lambda).$$

Then $\exists$ *a unitary operator* V *on* $\mathcal{S}_X$ *onto* $\mathcal{S}_Y$ *such that* $Y(\cdot)=V\cdot X(\cdot)$.

Proof. See Appendix B.

2.10 KERNEL THEOREM.† (Kolmogorov, Aronszajn, Pedrick). *Let* (i) W *be a Banach space over* $\mathbb{F}$ *and* W^* *be its adjoint*, (ii) $K(\cdot\,\cdot)$ *be a* PD *kernel on* $\Lambda\times\Lambda$ *to* $\mathrm{CL}(W, W^*)$. *Then*

(a) $\exists$ *a cardinal number* α *such that for all Hilbert spaces* $\mathfrak{H}$ *over* $\mathbb{F}$ *with* $\dim\mathfrak{H}\geqq\alpha$, $\exists$ *a function* $X(\cdot)$ *on* Λ *to* $\mathrm{CL}(W,\mathfrak{H})$ *for which the covariance kernel is* $K(\cdot\,\cdot)$, *i.e.*

$$\forall\lambda,\lambda'\in\Lambda,\qquad X(\lambda')^*X(\lambda)=K(\lambda,\lambda');$$

(b) $\forall\lambda,\lambda'\in\Lambda$ *and* $\forall w\in W$,

$$|X(\lambda)w|^2_{\mathfrak{H}}=\{K(\lambda,\lambda)(w)\}(w)\quad\textit{and}\quad |X(\lambda)|=\sqrt{|K(\lambda,\lambda)|}.$$

Proof. See Appendix C.

---

[13] For $A\subseteq\mathfrak{H}$, $\mathfrak{S}(A)\underset{d}{=}$ the (least closed linear) subspace spanned by A.

† See Note 1 added in proof.

2.11 *Remarks.* Thm. 2.10 subsumes a wide variety of known results. 1) When Λ is a singleton $\{\lambda\}$ and consequently $K(\lambda, \lambda)$ is just a nonnegative hermitian operator H on W to W^*, it reduces to a result due to Vakhania [33] and Chobanian [7] that H has a W-to-$\mathfrak{H}$ operator "square root" T such that $T^*T = H$. 2) When W is a separable Hilbert space, Λ a Hausdorff space and $K(\cdot\,\cdot)$ is continuous on $\Lambda \times \Lambda$, Thm. 2.10 yields (via Cor. 2.13 below) the recent result of Allen, Narcowich and Williams [1] cited earlier. 3) When $W = \mathbb{F}$, Thm. 2.10 reduces to the Aronszajn reproducing kernel theorem [3, p. 344] if we observe that to each $h \in \mathfrak{H}$ corresponds the $\mathbb{F}$-to-$\mathfrak{H}$ operator $T_h: c \to ch$, that $T_h^*(\cdot) = (\cdot, h)_{\mathfrak{H}}$ and consequently that $T_{h'}^* \cdot T_h: c \to (h, h')_{\mathfrak{H}}c$, $c \in \mathbb{F}$. 4) The same observation shows that when $W = \mathbb{F}$ and $\Lambda = \mathbb{N}_+$, Thm. 2.10 reduces to the original Kolmogorov Thm. 1.3.

2.12 COROLLARY. *Let* (i) *W be a Banach space and $\mathfrak{H}$ a Hilbert space, both over* $\mathbb{F}$, (ii) *$X(\cdot)$ be a function on Λ to* $\mathrm{CL}(W, \mathfrak{H})$ *having the covariance kernel $K(\cdot\,\cdot)$. Then $\forall \lambda, \lambda' \in \Lambda$ and $\forall w \in W$,*

(a) $|X(\lambda)(w) - X(\lambda')(w)|_{\mathfrak{H}}^2 = [\{K(\lambda, \lambda) + K(\lambda', \lambda') - K(\lambda, \lambda') - K(\lambda', \lambda)\}(w)](w)$,

(b) $|X(\lambda) - X(\lambda')|_B = \sqrt{|K(\lambda, \lambda) + K(\lambda', \lambda') - K(\lambda, \lambda') - K(\lambda', \lambda)|_B}$, *where the subscript B indicates the Banach norm.*

Proof. See Appendix D.

2.13 COROLLARY. *Let Λ be a Hausdorff space, and $W, \mathfrak{H}, X(\cdot), K(\cdot\,\cdot)$ be as in the last Corollary.*

(a) *If $K(\cdot\,\cdot)$ is continuous on $\Lambda \times \Lambda$ to the Banach space* $\mathrm{CL}(W, W^*)$, *then $X(\cdot)$ is continuous on Λ to the Banach space* $\mathrm{CL}(W, \mathfrak{H})$.

(b) *If $K(\cdot\,\cdot)$ is strongly continuous on $\Lambda \times \Lambda$ to* $\mathrm{CL}(W, W^*)$, i.e.

$$\forall \lambda_0, \lambda_0' \in \Lambda, \qquad \operatorname*{slim}_{(\lambda, \lambda') \to (\lambda_0, \lambda_0)} K(\lambda, \lambda') = K(\lambda_0, \lambda_0'),$$

then $X(\cdot)$ is strongly continuous on Λ to $\mathrm{CL}(W, \mathfrak{H})$.

Proof. See Appendix E.

Our next corollary reveals how the measurability of the kernel $K(\cdot\,\cdot)$ affects that of the variety $X(\cdot)$. It is convenient to introduce the following notation:

2.14 *Notation.* (a) For $\varnothing \neq \mathcal{U} \subseteq 2^\Lambda$ and $\varnothing \neq \mathcal{B} \subseteq 2^{\mathcal{X}}$,

$$\mathcal{M}(\mathcal{U}, \mathcal{B}) \underset{d}{=} \{f: f \in \mathcal{X}^\Lambda \text{ and } \forall B \in \mathcal{B}, f^{-1}(B) \in \mathcal{U}\},$$

i.e. $\mathcal{M}(\mathcal{U}, \mathcal{B})$ is the set of all $\mathcal{U}, \mathcal{B}$ *measurable* functions on Λ to $\mathcal{X}$.

(b) For any topological space $(\mathcal{X}, \tau)$ we shall write

$$\mathrm{Bl}(\mathcal{X}) \underset{d}{=} \sigma\text{-alg}(\tau) \underset{d}{=} \text{the } \sigma\text{-algebra over } \mathcal{X} \text{ generated by the topology } \tau;$$

this is the so-called *Borel algebra* over $\mathcal{X}$.

2.15 COROLLARY. *Let* (i) *$\mathcal{U}$ be a σ-algebra over Λ, and $W, \mathfrak{H}, X(\cdot), K(\cdot\,\cdot)$ be as in Corollary* 2.12; (ii) $\forall \lambda' \in \Lambda$, $K^{\lambda'}(\cdot) \underset{d}{=} K(\cdot, \lambda') \in \mathcal{M}(\mathcal{U}, \sigma\text{-alg}(\tau_s))$, *where τ_s is the strong operator topology for* $\mathrm{CL}(W, W^*)$. *Then*

(a) *$X(\cdot)$ is weakly $\mathcal{U}$-measurable on Λ to* $\mathrm{CL}(W, \mathfrak{H})$ *in the sense of* [12, p. 74, 3.5.5.(3)]; *more fully, $\forall w \in W$, $X(\cdot)w$ is $\mathcal{U}$-scalarly measurable on Λ to $\mathfrak{H}$, i.e.*

$$\forall w \in W \text{ and } \forall x \in \mathfrak{H}, \qquad (X(\cdot)w, x)_{\mathfrak{H}} \in \mathcal{M}(\mathcal{U}, \mathrm{Bl}(\mathbb{F}));$$

(b) *when $\mathfrak{H}$ is separable, $X(\cdot)$ is $\mathcal{U}$-strongly measurable on Λ to* $\mathrm{CL}(W, \mathfrak{H})$, *i.e.*

$$\forall w \in W, \qquad X(\cdot)(w) \in \mathcal{M}(\mathcal{U}, \mathrm{Bl}(\mathfrak{H})).$$

Proof. See Appendix F.

From topological and measurable parameter spaces Λ we turn to vectorial Λ. Our next result asserts that if Λ is a vector space over $\mathbb{F}$, and the PD kernel $K(\cdot\,\cdot)$ on $\Lambda\times\Lambda$ to $\mathrm{CL}(W, W^*)$ is sesquilinear, then the variety $X(\cdot)$ on Λ to $\mathrm{CL}(W, \mathfrak{H})$ with covariance kernel $K(\cdot\,\cdot)$ is a linear operator on Λ to $\mathrm{CL}(W, \mathfrak{H})$, and that for normed Λ and Lipschitzian $K(\cdot\,\cdot)$, the operator $X(\cdot)$ is continuous.

2.16 COROLLARY. *Let* (i) Λ *be a vector space over* $\mathbb{F}$, (ii) $K(\cdot\,\cdot)$ *be a* PD *kernel on* $\Lambda\times\Lambda$ *to* $\mathrm{CL}(W, W^*)$, *which is sesquilinear on* $\Lambda\times\Lambda$, (iii) $X(\cdot)$ *on* Λ *to* $\mathrm{CL}(W, \mathfrak{H})$ *be the variety with covariance kernel* $K(\cdot\,\cdot)$, $\mathfrak{H}$ *being a Hilbert space over* $\mathbb{F}$. *Then*

(a) $X(\cdot)$ *is a linear operator on* Λ *to* $\mathrm{CL}(W, \mathfrak{H})$;

(b) *when the vector space* Λ *is normed and the kernel* $K(\cdot\,\cdot)$ *is Lipschitzian, i.e.*

$$|K| \underset{d}{=} \sup_{0\neq\lambda\in\Lambda} \frac{|K(\lambda,\lambda)|}{|\lambda|^2} < \infty,$$

we have $X(\cdot)\in \mathrm{CL}(\Lambda, \mathrm{CL}(W, \mathfrak{H}))$ *and* $|X| = \sqrt{|K|}$.

Proof. See Appendix G.

3. Propagators on semi-groups. Hilbertian varieties $X(\cdot)$ parametrized over groups or semi-groups Λ have been studied extensively during the last few decades in connection with random processes and fields. It has been found that many such varieties carry with them, quite naturally, certain operator-valued functions $S(\cdot)$ on Λ, which act as their propagators or controllers,[14] and that the nature of $S(\cdot)$ is determined by the covariance kernel $K(\cdot\,\cdot)$ of $X(\cdot)$ as defined in 2.8(c).

An important case in point is that of a *stationary variety* $X(\cdot)$, i.e. one for which the covariance kernel $K(\cdot\,\cdot)$ on $\Lambda\times\Lambda$ is *translation invariant*:

$$\forall \lambda, \lambda', t\in\Lambda, \qquad K(\lambda+t, \lambda'+t) = K(\lambda, \lambda'). \tag{3.1}$$

The first theorem on the propagator of such a variety, due to Kolmogorov [15, § 2], asserts that for a stationary bisequence $(x_n)_{-\infty}^{\infty}$ in $\mathfrak{H}$, there exists a unitary operator U on $\mathscr{S}_x$ onto $\mathscr{S}_x$, where $\mathscr{S}_x$ is the subspace of $(x_n)_{-\infty}^{\infty}$, cf. 2.8(b), such that $\forall m, n\in \mathbb{N}$, $U^m(x_n) = x_{m+n}$. This U is called the *shift* of the bisequence $(x_n)_{-\infty}^{\infty}$, and $(U^n\colon n\in\mathbb{N})$ is called its *shift group*. The corresponding result for stationary curves and other stationary varieties is due to Karhunen [13, p. 55] and others. For varieties $X(\cdot)$ which are not themselves stationary in the sense of (3.1), but for which some easily associated variety $Y(\cdot)$ is stationary, such as helices in $\mathfrak{H}$, stationary $\mathfrak{H}$-valued orthogonally scattered measures and stationary W-to-$\mathfrak{H}$ quasi-isometric measures, the corresponding theorems are due to Schoenberg and von Neumann [34, p. 238] and the writer [20, p. 94], [21, p. 494]. The view that whenever some form of stationarity is inherent in a Hilbertian variety, it will possess isometric or unitary propagators is widely held, but no general theorem to this effect has been formulated. In this section we shall consider a wide concept of stationarity, 3.2(a), and state such a general theorem, Cor. 3.5, which will subsume all earlier results.

For unrestrictedly nonstationary varieties the propagator has not received much attention, and the only results we know are Getoor's on conditions for a sequence in $\mathfrak{H}$ to have a normal (or rather subnormal) shift operator [10]. They suggest that general semi-group versions of Halmos' theorem [11] on normal extensions can be viewed as results on subnormal propagators of general Hilbertian varieties.

[14] In those instances in which the members of Λ do not represent time or space-time or phase space, the term *controller* is more appropriate than the term *propagator*, cf. (5.11) et seq.

In order to cover all known instances of propagators and controllers we must include Hilbertian varieties $X(\cdot)$ over a domain Λ, which may or may not have a group structure, but which can be acted upon by a semi-group or group Γ. An instance is where Λ is the family of all subsets L of $\mathbb{R}$ of positive Lebesgue measure and $\Gamma=\mathbb{R}_{0+}$, and the action of t in $\mathbb{R}_{0+}$ transforms L into

$$t\oplus L \underset{d}{=} \{t+l: l\in L\}\in\Lambda.$$

Another instance with $\Gamma=\mathbb{R}_{0+}$ is where Λ is a solid body spinning about an axis with constant angular velocity, and $t\oplus\lambda$ is the position of the particle λ in Λ t seconds later. Generally, $t\oplus\lambda$ is $T_t(\lambda)$, where $(T_t: t\in\Gamma)$ is a family of transformations on Λ to Λ parametrized on the semi-group Γ. $T_t(\lambda)$ may be regarded as the position at instant t of a particle which was at λ at instant 0. In these instances $\Lambda\neq\Gamma$. On the other hand as indicated earlier, we will have $\Lambda=\Gamma$ in many cases. The concepts of action, propagator or controller, and stationarity, which cover all cases, are defined as follows:

3.2 Definition. Let (i) $X(\cdot)$ be a function on Λ to $\mathrm{CL}(W,\mathfrak{H})$, where Λ is any set, W is a Banach space over $\mathbb{F}$ and $\mathfrak{H}$ a Hilbert space over $\mathbb{F}$, (ii) Γ be a semi-group under $+$ (not necessarily Abelian) with neutral element 0. Then

(a) we say that Γ *acts on* Λ iff $\exists$ a binary operation $\oplus$ on $\Gamma\times\Lambda$ to Λ such that

$$\forall s,t\in\Gamma \text{ and } \forall\lambda\in\Lambda,\qquad (s+t)\oplus\lambda=s\oplus(t\oplus\lambda)\quad\text{and}\quad 0\oplus\lambda=\lambda.$$

(b) $S(\cdot)$ is called a *propagator* or *controller of* $X(\cdot)$, iff $S(\cdot)$ is a function on Γ to $\mathrm{CL}(\mathscr{S}_X,\mathscr{S}_X)$, where $\mathscr{S}_X$ is as in 2.8(b), and

$$\forall t\in\Gamma \text{ and } \forall\lambda\in\Lambda,\qquad S(t)\circ X(\lambda)=X(t\oplus\lambda).$$

(c) $X(\cdot)$ is called *stationary* (*in the wide sense*), iff its covariance kernel $K(\cdot\,\cdot)$ on $\Lambda\times\Lambda$, cf. 2.8(c), is translation-invariant in the sense that

$$\forall t\in\Gamma \text{ and } \forall\lambda,\lambda'\in\Lambda,\qquad K(t\oplus\lambda, t\oplus\lambda')=K(\lambda,\lambda').$$

The case in which Λ is itself a semi-group is of course subsumed in the last definition by just letting $\oplus$ be $+$.

3.3 Proposition. *Let* (i) $\Lambda,\Gamma,W,\mathfrak{H},X(\cdot)$ *be as in* 3.2(i), (ii); (ii) *the semi-group* Γ *act on* Λ; (iii) $X(\cdot)$ *possess a propagator or controller* $S(\cdot)$. *Then*

(a) $(S(t)\colon t\in\Gamma)$ *is a semi-group of operators in* $\mathrm{CL}(\mathscr{S}_X,\mathscr{S}_X)$, *i.e.*

$$S(0)=I_{\mathscr{S}_X},\quad\text{and}\quad \forall s,t\in\Gamma,\qquad S(s+t)=S(s)\cdot S(t);$$

(b) *when* Γ *is a group, we have* $(S(t)\colon t\in\Gamma)$ *is a group of one-one and onto operators in* $\mathrm{CL}(\mathscr{S}_X,\mathscr{S}_X)$.

Proof. (a) By (iii),

(1) $$\forall t\in\Gamma,\qquad S(t)\in\mathrm{CL}(\mathscr{S}_X,\mathscr{S}_X).$$

Now let

(2) $$\mathscr{D}\underset{d}{=}\bigcup_{\lambda\in\Lambda}X(\lambda)(W),$$

(3) $\langle\mathscr{D}\rangle$ be the linear manifold spanned by $\mathscr{D}$ in $\mathfrak{H}$.

(4) $$\begin{cases}\forall t\in\Gamma,\qquad S_t=\mathrm{Rstr.}_{\mathscr{D}}S(t),\\ \langle S_t\rangle\underset{d}{=}\text{the linear manifold spanned by } S_t \text{ in } \mathfrak{H}\times\mathfrak{H}.\end{cases}$$

Then obviously

(5) $\langle S_t\rangle \underset{d}{=} \mathrm{Rstr.}_{\langle\mathscr{D}\rangle}S(t)$ and $S(t)=\mathrm{cls.}\langle S_t\rangle \underset{d}{=}$ the closure of $\langle S_t\rangle$.

Now by (ii) and (iii)

(6) $\forall t\in\Gamma$ and $\forall\lambda\in\Lambda,\qquad S(t)\circ X(\lambda)=X(t\oplus\lambda)$.

From this we easily infer that

$$S(s+t)\circ X(\lambda)=S(s)\circ S(t)\circ X(\lambda),$$

and therefore that $S_{s+t}=S_s\circ S_t$. From this it follows in turn, since Range $S_t\subseteq\mathscr{D}=$ domain of S_s, and $\langle S_s\rangle$ is single-valued, that $\langle S_{s+t}\rangle=\langle S_s\rangle\circ\langle S_t\rangle$, and hence from (5) that

(7) $$S(s+t)=S(s)\circ S(t).$$

This last hinges on the result that if T_1, T_2 are continuous linear operators and the domain of T_2 contains the range of T_1, then cls. $(T_2T_1)=$ cls. $T_2\circ$ cls. T_1.

Finally, taking $t=0$ in (6), we get $S(0)\circ X(\lambda)=X(\lambda)$, which shows that $S_0=I_{\mathscr{D}}$, $\langle S_0\rangle=I_{\langle\mathscr{D}\rangle}$ and therefore

(8) $$S(0)=I_{\mathscr{S}_X}.$$

By (7), (8) we have (a).

(b) When Γ is a group, we have $\forall t\in\Gamma$,

$$S(-t+t)=S(0)=S(t+(-t)),$$

i.e. by (7) and (8)

$$S(-t)\circ S(t)=I_{\mathscr{S}_X}=S(t)\circ S(-t).$$

This shows that $S(t)$ is one-one on $\mathscr{S}_X$ onto $\mathscr{S}_X$, and that $S(t)^{-1}=S(-t)$. Thus (b). □

Given a Hilbertian variety $X(\cdot)$ parametrized as in 3.2, the first question we must answer pertains to the conditions which $X(\cdot)$ must satisfy in order that it may possess a propagator $S(\cdot)$. We may then ask for the extra conditions that $X(\cdot)$ must satisfy in order that $S(\cdot)$ may be of a specific type. The following general theorem gives a full answer to the first question. In essence it is a completed, operator-extension for a not necessarily Abelian semi-group Γ and a Banach space W of two lemmas of Getoor [10, 2.1, 2.2], and our proof is an extension of his own.

3.4 THEOREM (existence of propagator). *Let* Λ, Γ, W, $\mathfrak{H}$, $X(\cdot)$ *be as in* 3.2(i), (ii), *and the semi-group* Γ *acts on* Λ. *Then*

(a) *the following conditions are equivalent*:

(α) $X(\cdot)$ *has a propagator* $S(\cdot)$ *on* Γ,

(β) $X(\cdot)$ *satisfies the following translational inequality*: $\exists$ *a function* $\beta(\cdot)$ *on* Γ *to* $\mathbb{R}_{0+}\ni\forall$ *functions* $w(\cdot)$ *on* Λ *to* W, $\forall r\in\mathbb{N}_+$, $\forall\lambda_1,\cdots,\lambda_r\in\Lambda$ *and* $\forall t\in\Gamma$,

$$\left|\sum_{i=1}^r X(t\oplus\lambda_i)w(\lambda_i)\right|_{\mathfrak{H}}\leqq\beta(t)\left|\sum_{i=1}^r X(\lambda_i)w(\lambda_i)\right|_{\mathfrak{H}},$$

(γ) *the covariance kernel* $K(\cdot\,\cdot)$ *of* $X(\cdot)$, *cf.* 2.8(c), *satisfies the following translational inequality*: $\exists$ *a function* $\gamma(\cdot)$ *on* Γ *to* $\mathbb{R}_{0+}\ni\forall$ *functions* $w(\cdot)$ *on* Λ *to* W, $\forall r\in\mathbb{N}_+$, $\forall\lambda_1,\cdots,\lambda_r\in\Lambda$, *and* $\forall t\in\Gamma$,

$$0\leqq\sum_{i=1}^r\sum_{j=1}^r[\{K(t\oplus\lambda_i,t\oplus\lambda_j)\}w(\lambda_i)]w(\lambda_j)\leqq\gamma(t)\sum_{i=1}^r\sum_{j=1}^r[\{K(\lambda_i,\lambda_j)\}w(\lambda_i)]w(\lambda_j);$$

(b) *when* (α), (β) *or* (γ) *hold, the function* $S(\cdot)$ *is unique, and* $|S(t)|_B = \beta(t) = \sqrt{\gamma(t)}$ *with the best choice of* $\beta(\cdot)$, $\gamma(\cdot)$, *the subscript B denoting the Banach norm.*

Proof. (a) We adopt the abbreviation $w_i = w(\lambda_i)$. Then the equivalence of (β) and (γ) follows easily from the equality, cf. 2.2(c),

$$(1)\qquad \left|\sum_{i=1}^{r} X(\lambda_i)w_i\right|_{\mathfrak{H}}^2 = \sum_{i=1}^{r}\sum_{j=1}^{r} (X(\lambda_i)w_i, X(\lambda_j)w_j)_{\mathfrak{H}} = \sum_{i=1}^{r}\sum_{j=1}^{r} [\{K(\lambda_i, \lambda_j)\}(w_i)](w_j),$$

and the corresponding equality in which $t\oplus\lambda_i$ replaces λ_i for $i = 1, \cdots, r$. Hence it only remains to prove that (α)⇔(β).

Let (α) hold. Then obviously so does (β) with the function $\beta(\cdot) \underset{d}{=} |S(\cdot)|_B$ on Γ, since by Def. 3.2(b), $\forall t \in \Gamma$

$$\left|\sum_{i=1}^{r} X(t\oplus\lambda_i)w_i\right|_{\mathfrak{H}} = \left|\sum_{i=1}^{r} S(t)\cdot X(\lambda_i)w_i\right|_{\mathfrak{H}} = \left|S(t)\left\{\sum_{i=1}^{r} X(\lambda_i)w_i\right\}\right|_{\mathfrak{H}} \leqq \beta(t)\left|\sum_{i=1}^{r} X(\lambda_i)w_i\right|_{\mathfrak{H}}.$$

Next let (β) hold. Fix $t\in\Gamma$ and let

$$(2)\qquad \mathscr{D} \underset{d}{=} \bigcup_{\lambda\in\Lambda} X(\lambda)(W),$$

(3) $\langle\mathscr{D}\rangle$ be the linear manifold spanned by $\mathscr{D}$ in $\mathfrak{H}$,

(4) $S_t \underset{d}{=} \{(X(\lambda)w; X(t\oplus\lambda)w): \lambda\in\Lambda \text{ and } w\in W\} \subseteq \mathfrak{H}\times\mathfrak{H}$,

(5) $\langle S_t\rangle$ be the linear manifold spanned by S_t in $\mathfrak{H}\times\mathfrak{H}$.

Then $\langle S_t\rangle$ is a linear relation on $\langle\mathscr{D}\rangle$ to $\langle\mathscr{D}\rangle$; in fact

$$(6)\quad \langle S_t\rangle = \left\{\left(\sum_{i=1}^{r} X(\lambda_i)w_i; \sum_{i=1}^{r} X(t\oplus\lambda_i)w_i\right): r\in\mathbb{N}_+, \lambda_1, \cdots, \lambda_r\in\Lambda \text{ and } w_i, \cdots, w_r \in W\right\}.$$

Since by (β),

$$\sum_{i=1}^{r} X(\lambda_i)w_i = 0 \Rightarrow \sum_{i=1}^{r} X(t\oplus\lambda_i)w_i = 0,$$

it follows from (6) that $\langle S_t\rangle$ is single-valued at 0, and therefore throughout its domain $\langle\mathscr{D}\rangle$. Moreover, if $x\in\langle\mathscr{D}\rangle$, say $x \underset{d}{=} \sum_{i=1}^{r} X(\lambda_i)w_i$, then by (6), the single-valuedness of S_t and (β),

$$|\langle S_t\rangle x|_{\mathfrak{H}} = \left|\sum_{i=1}^{r} X(t\oplus\lambda_i)w_i\right|_{\mathfrak{H}} \leqq \beta(t)|x|.$$

We have thus shown that

$$(7)\qquad \langle S_t\rangle \in \mathrm{CL}(\langle\mathscr{D}\rangle, \langle\mathscr{D}\rangle) \quad\text{and}\quad |\langle S_t\rangle|_B \leqq \beta(t).$$

Since cls. $\langle\mathscr{D}\rangle = \mathscr{S}_X$, it follows at once from (7) and the extension principle for continuous linear operators that

$$(8)\qquad S(t) \underset{d}{=} \text{cls.}\,\langle S_t\rangle \in \mathrm{CL}(\mathscr{S}_X, \mathscr{S}_X) \quad\text{and}\quad |S(t)|_B \leqq \beta(t).$$

Finally, since by (4), $\forall w \in W$,

$$\{S(t) \circ X(\lambda)\}w = S_t\{X(\lambda)w\} = X(t \oplus \lambda)w,$$

we have of course

$$(9) \qquad \forall t \in \Gamma, \qquad S(t) \circ X(\lambda) = X(t \oplus \lambda).$$

Thus $S(\cdot)$ is the propagator of $X(\cdot)$, i.e. we have (α).

This completes the proof of (a).

(b) Were $T(\cdot)$ on Γ to $\mathrm{CL}(\mathscr{S}_X, \mathscr{S}_X)$ another propagator of $X(\cdot)$, we would have by Def. 3.2(b),

$$T(t) \circ X(\lambda) = X(t \oplus \lambda) = S(t) \circ X(\lambda),$$

and therefore $T(t) = S(t)$ on $\mathscr{D}$. It would follow that $T(t) = S(t)$ on cls. $\langle \mathscr{D} \rangle$, i.e. on $\mathscr{S}_X$. Thus the propagator is unique. Also, we can take $\beta(\cdot) = |S(\cdot)|_B$, cf. paragraph 2 of this proof, and take $\gamma(\cdot) = \beta(\cdot)^2$ in view of (1) and of (β)⇔(γ). Thus we have (b). □

An easy consequence of the last theorem is the following corollary on stationary varieties in the wide sense of Def. 3.2(c), which subsumes all earlier results of its kind; cf. second paragraph of this section.

3.5 Corollary (stationary varieties). *Let* Λ, Γ, W, $\mathfrak{H}$, $X(\cdot)$ *be as in* 3.2(i), (ii). *Then*

(a) *the following conditions are equivalent*:

(α) $X(\cdot)$ *is stationary*,

(β) $X(\cdot)$ *has a propagator semi-group* $S(\cdot)$ *on* Γ *such that* $\forall t \in \Gamma$, $S(t)$ *is an isometry on* $\mathscr{S}_X$ *into* $\mathscr{S}_X$;

(b) *when* (α) *or* (β) *hold, the semi-group* $S(\cdot)$ *is unique*;

(c) *when* Γ *is a group and* (α) *or* (β) *hold*, $S(\cdot)$ *is a group of unitary operators on* $\mathscr{S}_X$ *onto* $\mathscr{S}_X$.

Proof. (a) Let (β) hold. Then obviously $\forall t \in \Gamma$, $S(t)^*S(t) = I_{\mathscr{S}_X}$, whence $\forall \lambda, \lambda' \in \Lambda$

$$\begin{aligned} K(t \oplus \lambda, t \oplus \lambda') &\underset{d}{=} X(t \oplus \lambda')^* X(t \oplus \lambda) \\ &= X(\lambda')^* S(t)^* S(t) X(\lambda) \\ &= X(\lambda')^* X(\lambda) \underset{d}{=} K(\lambda, \lambda'); \end{aligned}$$

i.e. we have (α); cf. Def. 3.2(c).

Next let (α) hold. Then the condition 3.4(γ) certainly holds with $\gamma(t) = 1$, since

$$(1) \qquad 0 \leqq \sum_{i=1}^{r} \sum_{j=1}^{r} [\{K(t \oplus \lambda_i, t \oplus \lambda_j)\}(w_i)](w_j) = \sum_{i=1}^{r} \sum_{j=1}^{r} [\{K(\lambda_i, \lambda_j)\}(w_i)](w_j).$$

Hence by Thm. 3.4, the variety $X(\cdot)$ possesses a propagator semi-group $S(\cdot)$ on Γ. To complete the proof of (β) it only remains to show that

$$(\mathrm{A}) \qquad \forall t \in \Gamma, \qquad S(t) \text{ is an isometry on } \mathscr{S}_X \text{ into } \mathscr{S}_X.$$

Proof of (A). First let $x \in \langle \mathscr{D} \rangle$, where $\mathscr{D} \underset{d}{=} \bigcup_{\lambda \in \Lambda} X(\lambda)(W)$; say $x = \sum_{i=1}^{r} X(\lambda_i)(w_i)$.

Then as in the proof of Thm. 3.4, $\forall t \in \Lambda$

$$|S(t)x|_{\mathfrak{H}}^2 = \left|\sum_{i=1}^r X(t\oplus\lambda_i)w_i\right|_{\mathfrak{H}}^2$$

$$= \sum_{i=1}^r \sum_{j=1}^r [\{K(t\oplus\lambda_i, t\oplus\lambda_j)\}(w_i)](w_j)$$

$$= \sum_{i=1}^r \sum_{j=1}^r [\{K(\lambda_i, \lambda_j)\}(w_i)](w_j)$$

$$= \left|\sum_{i=1}^r X(\lambda_i)w_i\right|_{\mathfrak{H}}^2 = |x|_{\mathfrak{H}}^2.$$

Thus $\text{Rstr.}_{\langle\mathscr{D}\rangle}S(t)$ is an isometry on $\langle\mathscr{D}\rangle$ to $\langle\mathscr{D}\rangle$. From this it follows readily, since cls. $\langle\mathscr{D}\rangle = \mathscr{S}_X$, that $S(t)$ is an isometry on $\mathscr{S}_X$ to $\mathscr{S}_X$. Thus (A) is proved. This completes the proof of (a).

Parts (b) and (c) follow at once from (a) and 3.4(b) and 3.3(b). □

When the semi-group Γ, which acts on the parameter space Λ via the operation $\oplus$, is topological, the concept of a continuous kernel $K(\cdot\,\cdot)$ on $\Lambda\times\Lambda$ makes sense. It is natural to ask if the propagator $S(\cdot)$ of a variety $X(\cdot)$ on Λ having such a covariance kernel $K(\cdot\,\cdot)$ will itself be continuous. The answer is affirmative as the following corollary shows:

3.6 COROLLARY (strongly continuous propagator). *Let*

(i) $\Lambda, \Gamma, W, \mathfrak{H}, X(\cdot)$ *be as in* 3.2(i), (ii);

(ii) *the semi-group* Γ *in* (i) *be topological, i.e.* Γ *be a Hausdorff space and the operation* $+$ *be continuous on* $\Gamma\times\Gamma$ *to* Γ *with respect to the topology of* Γ *and the corresponding weak product topology for* $\Gamma\times\Gamma$;

(iii) *the covariance kernel* $K(\cdot\,\cdot)$ *of* $X(\cdot)$ *be strongly continuous on* $\Lambda\times\Lambda$ *to* $\text{CL}(W, W^*)$ *in the sense that*

$$\forall\lambda, \lambda' \in \Lambda, \qquad \operatorname*{slim}_{(t,t')\to(0,0)} K(t\oplus\lambda, t'\oplus\lambda') = K(\lambda, \lambda');$$

(iv) $\exists$ *a function* $\beta(\cdot)$ *on* Γ *to* $\mathbb{R}_{0+}$ *which satisfies the condition* 3.4(β) *and is bounded on a neighborhood* V_0 *of* 0 *in* Γ.

Then $S(\cdot)$ *is strongly continuous on* Λ *to* $\text{CL}(\mathscr{S}_X, \mathscr{S}_X)$.

Proof. Grant temporarily that

(I) $$\operatorname*{slim}_{s\to 0} S(s) = I_{\mathscr{S}_X}.$$

Then, cf. 3.2(b), $\forall x \in \mathscr{S}_X$ and $\forall s, t \in \Lambda$,

$$|S(t+s)x - S(t)x|_{\mathfrak{H}} = |S(t)\{S(s)-I\}x|_{\mathfrak{H}} \leqq |S(t)|_B \cdot |S(s)x - x|_{\mathfrak{H}}.$$

Since by (I) the RHS $\to 0$ as $s\to 0$, we have $\text{slim}_{s\to 0} S(t+s) = S(t)$ as desired. Hence it only remains to prove (I).

Proof of (I). Let $s\in\Gamma$ and $x \in \mathscr{D} \underset{d}{=} \bigcup_{\lambda\in\Lambda} X(\lambda)(W)$, say $x = X(\lambda)(w)$. Then by 3.2(b) and 2.12(a)

$$|S(s)x - x|_{\mathfrak{H}}^2 = |S(s)X(\lambda)w - X(\lambda)w|_{\mathfrak{H}}^2 = |X(s\oplus\lambda)w - X(\lambda)w|_{\mathfrak{H}}^2$$

$$= [\{K(s\oplus\lambda, s\oplus\lambda) - K(s\oplus\lambda, \lambda) - K(\lambda, s\oplus\lambda) + K(\lambda,\lambda)\}(w)](w)$$

(1) $$\leqq |\{K(s\oplus\lambda, s\oplus\lambda) - K(s\oplus\lambda, \lambda) - K(\lambda, s\oplus\lambda) + K(\lambda,\lambda)\}(w)|_{W^*}|w|.$$

Now by (iii), the first factor on RHS(1) tends to 0 as $s\to 0$. Hence by (1), $\lim_{s\to 0} S(s)x = x$

in $\mathfrak{H}$. It follows from the linearity of the $S(s)$ that this limiting equality holds for all finite linear combinations of x in $\mathscr{D}$, i.e. we have

$$(2) \qquad \langle \mathscr{D} \rangle \subseteq \text{Dom.} \operatorname*{slim}_{s \to 0} S(s) \quad \text{and} \quad \text{Rstr.} \langle \mathscr{D} \rangle \operatorname*{slim}_{s \to 0} S(s) = I_{\langle \mathscr{D} \rangle}.$$

Now by Thm. 3.4(b) and (iv)

$$\forall s \in V_0, \qquad |S(s)|_B = \beta(s) \leqq m_0, \quad \text{say.}$$

We conclude from the lemma on strong convergence, just following this proof, that Dom. $\operatorname{slim}_{s\to 0} S(s)$ is a closed subspace, and $\operatorname{slim}_{s\to 0} S(s)$ is a continuous linear operator. Since cls. $\langle \mathscr{D} \rangle = \mathscr{S}_X$, it follows from (2) that the domain is $\mathscr{S}_X$; thus

$$\operatorname*{slim}_{s \to 0} S(s) \in \text{CL}(\mathscr{S}_X, \mathscr{S}_X).$$

Moreover, from the second half of (2) and the principle of uniqueness of continuous extension, $\operatorname{slim}_{s\to 0} S(s) = I_{\mathscr{S}_X}$. Thus (I) is established and the proof is over. □

In the last proof we have appealed to the following lemma, the proof of which we leave to the reader.

LEMMA. *Let* (i) Γ *be a topological space,* (ii) $\mathscr{X}$ *be a Banach space,* (iii) $S(\cdot)$ *a function on* Γ *to* $\text{CL}(\mathscr{X}, \mathscr{X})$, (iv) $a \in \Gamma$, *and* $\exists$ *a neighbourhood* V_a *of* a $\ni \sup_{t \in V_a} |S(t)|_B < \infty$, (v) $L = \operatorname{slim}_{t \to a} S(t)$.
Then

$$\mathscr{D}_L \underset{d}{=} \textit{Domain of } L \textit{ is a closed subspace of } \mathscr{X},$$

and

$$L \in \text{CL}(\mathscr{D}_L, \mathscr{X}) \quad \text{and} \quad |L|_B \leqq \varliminf_{t \to a} |S(t)|_B < \infty.$$

4. Propagators on involutory semi-groups. We now turn to the fruitful concepts of an involutory semi-group and of PD functions thereon due to Nagy [26, p. 20, 21].

4.1 DEFINITION. (a) Γ is called an *involutory semi-group* (briefly, *s.g.) iff (i) Γ is an (additive, not necessary Abelian) s.g. with neutral element 0, (ii) Γ possesses an involution $(\cdot)^*$ satisfying

$$\forall s, t \in \Gamma, \quad (t^*)^* = t, \quad (s+t)^* = t^* + s^*, \quad 0^* = 0.$$

(b) A function $R(\cdot)$ on a *s.g. Γ to $\text{CL}(W, W^*)$, where W is a Banach space, is called *positive definite* (PD) iff the kernel $K(\cdot\,\cdot)$ defined by

$$\forall s, t \in \Gamma, \qquad K(s,t) = R(t^* + s)$$

is PD on $\Gamma \times \Gamma$ to $\text{CL}(W, W^*)$ in the sense of 2.5.[15]

Adopting the terminology of operator algebras, we define

$$(4.2) \qquad \begin{cases} \Gamma_{(n)} \underset{d}{=} \{t : t \in \Gamma \quad \text{and} \quad t^* + t = t + t^*\}, \\ \Gamma_{sa} = \{t : t \in \Gamma \quad \text{and} \quad t = t^*\}, \\ \Gamma_{0+} = \{t : t \in \Gamma \quad \text{and} \quad \exists s \in \Gamma \ni t = s^* + s\}, \end{cases}$$

and to refer to the members of $\Gamma_{(n)}$, Γ_{sa} and Γ_{0+} as the *normal, self-adjoint* and *nonnegative* (≥ 0) elements of Γ.

[15] In the special case in which $W = \mathbb{F}$, Γ is a group and $t^* = -t$; this definition reduces to the classical one.

Obviously,

$$(4.3)\qquad 0\in\Gamma_{0+}\subseteq\Gamma_{sa}\subseteq\Gamma_{(n)}\subseteq\Gamma,$$

but Γ_{sa} will be a (sub-) semi-group iff its elements commute. In particular, Γ_{sa} is a semi-group when Γ is Abelian. It is also easy to check that

$$(4.4)\qquad \begin{cases} t \text{ is normal } \Rightarrow \forall n\in\mathbb{N}_{0+},\ nt \text{ is normal},\\ t \text{ is sa } \Rightarrow \forall n\in\mathbb{N}_{0+},\ nt \text{ is sa},\\ t\geq 0 \Rightarrow \forall n\in\mathbb{N}_{0+},\ nt\geq 0.\end{cases}$$

Our major objective is to show that in the analogue of Thm. 3.4 for the action of an involutary semi-group, we can replace the translational inequality 3.4(γ) imposed on $K(\cdot\,\cdot)$ in order to get the propagator by the much milder requirements:

$$(4.5)\begin{cases}\text{(a)}\quad \forall\lambda,\lambda'\in\Lambda \text{ and } \forall t\in\Gamma,\ K(\lambda, t^*\oplus\lambda')=K(t\oplus\lambda,\lambda'),\\ \text{(b)}\quad \exists \text{ a function } \gamma(\cdot) \text{ on } \Gamma \text{ to } \mathbb{R}_{0+}\ni\forall\lambda\in\Lambda \text{ and } \forall t\in\Gamma,\\ \qquad\qquad 0\leq K(t\oplus\lambda, t\oplus\lambda)\leq\gamma(t)\cdot K(\lambda,\lambda).\end{cases}$$

We shall refer to these as the *transfer property* and the *mild translational property*.

The replacement of 3.4(γ) by (4.5) is a considerable simplification, since any kernel of the type described in 4.1(b) automatically satisfies (4.5(a), and the inequality

$$0\leqq K(t\oplus\lambda, t\oplus\lambda)(w)(w)\leqq\gamma(t)\cdot K(\lambda,\lambda)(w)(w),$$

which is an equivalent rendering of (4.5)(b), is much simpler than (3.4)(γ).

The genesis of (4.5) calls for explanation. Nagy's principal thm. [26, p. 21] concerning the dilation of a PD function $R(\cdot)$ on a *s.g. Γ contains a premise (c) affirming a translational inequality of the type 3.4(γ). Now in 1955 Bram [6, Thm. 1] proved the redundancy of a similar premise appearing in Halmo's theorem on normal extensions (cf. [26, p. 19]). This major advance raised the question of a possible mitigation if not elimination of Nagy's inequality (c) itself. Recently the writer learned from Dr. Arveson about a proof he had obtained in 1971, but not published, that the Nagy inequality is replaceable by the requirement that $R(\cdot)$ be bounded on Γ. This proof leaned heavily on Stinespring's work on C^*-algebras [32]. The writer in his own investigations of the single-valuedness and continuity of the propagator then found the conditions (4.5).

It will be understood in what follows that

(4.6)
- (i) *X is a function on Λ to* CL$(W,\mathfrak{H})$, *where Λ is any set, W is a Banach space over $\mathbb{F}$ and $\mathfrak{H}$ a Hilbert space over $\mathbb{F}$,*
- (ii) $\mathcal{D}=\bigcup_{\lambda\in\Lambda}X(\lambda)(W)$,
- (iii) *Γ is a* *s.g. (*not necessarily Abelian*) *With neutral element* 0 *that acts on* Λ (cf. Def. 3.2(a)).

We first assert that the transfer property (4.5)(a) alone yields single-valued, closed, but possibly discontinuous propagators on domains containing the e.d. linear manifold $\langle\mathcal{D}\rangle$. This result is an important step towards our goal, but has some interest of its own:

4.7 THEOREM (existence of closed, densely defined "propagator"). *Let* Λ, W, $\mathfrak{H}$, $X(\cdot)$, $\mathcal{D}$, Γ *be as in* (4.6). *Then the following conditions are equivalent*:

(α) $\forall t\in\Gamma$, $\exists a$ (s.v.) *closed linear operator $S(t)$ from $\mathcal{S}_X$ to $\mathcal{S}_X$ with domain* $\mathcal{D}_t\supseteq\langle\mathcal{D}\rangle$ *such that*

$$\forall\lambda\in\Lambda,\quad S(t)\cdot X(\lambda)=X(t\oplus\lambda)\quad \textit{and}\quad \text{Rstr.}_{\mathcal{D}}S(t^*)\subseteq S(t)^*,$$

(β) *the covariance kernel $K(\cdot\,\cdot)$ of $X(\cdot)$ possesses the transfer property* (4.5)(a).

Proof. Let (α) hold, and $t \in \Gamma$. First observe that

(1) $S(t)^*$ is a s.v. closed linear operator from $\mathcal{S}_X$ to $\mathcal{S}_X$ with an e.d. domain $\mathcal{D}_t^* \supseteq \langle \mathcal{D} \rangle$.

For $S(t)^*$ is always closed; $S(t)^*$ is s.v., since $\mathcal{D}_t$ is e.d.; $\mathcal{D}_t^*$ is e.d., since $S(t)$ is s.v.; and $\langle \mathcal{D} \rangle \subseteq \mathcal{D}_t^*$, since by (α) $\mathcal{D} \subseteq \mathcal{D}_t^*$.

Now let $\lambda' \in \Lambda$. Since by (α), $S(t^*) = S(t)^*$ on $\mathcal{D}$, therefore $X(\lambda')^* \cdot S(t^*) = X(\lambda')^* \cdot S(t)^*$ on $\mathcal{D}$. But the last operator is invariably a suboperator of $[S(t) \cdot X(\lambda')]^*$ which by (α) is $X(t \oplus \lambda')^*$. Thus

$$X(\lambda')^* \cdot S(t^*) = X(t \oplus \lambda')^* \quad \text{on } \mathcal{D}.$$

Replacing t by t^* and applying both operators to the vector $X(\lambda)w \in \mathcal{D}$, where $\lambda \in \Lambda$ and $w \in W$, we get

$$X(\lambda')^* \cdot S(t)X(\lambda)(w) = X(t^* \oplus \lambda')^* \cdot X(\lambda)(w).$$

But by (α), $S(t)X(\lambda) = X(t \oplus \lambda)$, and so

$$X(\lambda')^* \cdot X(t \oplus \lambda)(w) = X(t^* \oplus \lambda')^* \cdot X(\lambda)(w).$$

As this holds $\forall w \in W$, we have

$$X(\lambda')^* \cdot X(t + \lambda) = X(t^* \oplus \lambda')^* \cdot X(\lambda), \quad \text{i.e. (4.5)(a).}$$

Thus (β) holds.

Next let (β) hold. Fix $t \in \Gamma$, and define S_t, $\langle S_t \rangle$ as in (4), (5) of Proof of 3.4, and define

(1′) $$S(t) \underset{d}{=} \text{cls.} \langle S_t \rangle = \text{the closure of } \langle S_t \rangle \text{ in } \mathcal{S}_X \times \mathcal{S}_X.$$

Then

(2) $S(t)$ is a closed linear relation from $\mathcal{S}_X$ to $\mathcal{S}_X$.

We assert that

(I) $S(t)$ is single-valued.

Proof of (I). First observe that by the transfer property (4.5)(a),

(3) $$\forall t \in \Gamma \text{ and } \forall \lambda, \lambda' \in \Lambda, \qquad X(t^* \oplus \lambda')^* \cdot X(\lambda) = X(\lambda')^* \cdot X(t \oplus \lambda).$$

Now let the ordered pair $(0; y) \in S(t)$. Then $\forall n \in \mathbb{N}_+$, $\exists (x_n; y_n) \in \langle S_t \rangle$ such that

(4) $$x_n \to 0 \quad \text{and} \quad y_n \to y \quad \text{in } \mathcal{S}_X, \quad \text{as } n \to \infty.$$

Now with an obvious notation let (cf. (6) in Proof of 3.4)

$$x_n = \sum_{i=1}^{r_n} X(\lambda_i^n) w_i^n, \qquad y_n = \sum_{i=1}^{r_n} X(t \oplus \lambda_i^n) w_i^n.$$

Then from the linearity of $X(t^* \oplus \lambda)^*$, $X(\lambda)^*$ and equation (3),

$$\forall \lambda \in \Lambda, \qquad X(t^* \oplus \lambda)^*(x_n) = X(\lambda)^*(y_n).$$

From (4) and the continuity of the operators $X(t^* \oplus \lambda)^*$, $X(\lambda)^*$ on $\mathcal{S}_X$, we get on letting $n \to \infty$,

$$0 = X(t^* \oplus \lambda)^*(0) = X(\lambda)^*(y).$$

Thus y is in the null space of $X(\lambda)^*$, i.e., cf. 2.2(c), $\forall \lambda \in \Lambda$, $y \perp X(\lambda)(W)$. It follows that

$$y \perp \mathfrak{S}\{X(\lambda)(W) \colon \lambda \in \Lambda\} = \mathcal{S}_X.$$

But since $y_n \in \langle \mathscr{D} \rangle \subseteq \mathscr{S}_X$, therefore by (4), $y \in \mathscr{S}_X$. Thus $y = 0$. This means that $S(t)$ is single-valued at 0, and, being linear, it is single-valued throughout its domain. Thus (I) is proved.

Combining (2) and (I) we have

(5) $S(t)$ is a (s.v.) closed linear operator from $\mathscr{S}_X$ to $\mathscr{S}_X$ with domain $\mathscr{D}_t \supseteq \langle \mathscr{D} \rangle$.

Also, cf. (4) in proof of 3.4, $\forall \lambda \in \Lambda$ and $\forall w \in W$

$$\{S(t) \cdot X(\lambda)\}(w) = S_t\{X(\lambda)w\} = X(t \oplus \lambda)(w).$$

Thus

(6) $$\forall t \in \Gamma \text{ and } \forall \lambda \in \Lambda, \qquad S(t) \cdot X(\lambda) = X(t \oplus \lambda).$$

We now turn to the inclusion relation in (α), which alone remains unproved. Since the e.d. set $\langle \mathscr{D} \rangle$ is contained in $\mathscr{D}_t$, it follows that $S(t)^*$ is s.v., and by (I), $S(t)^*$ has a domain $\mathscr{D}_t^*$ e.d. in $\mathscr{S}_X$. Now grant for a moment that

(II) $$\forall x \in \mathscr{D}_t \text{ and } \forall x' \in \mathscr{D}, \qquad (S(t)x, x') = (x, S(t^*)x').$$

Then by the definition of the adjoint

$$\forall x' \in \mathscr{D}, \quad x' \in \mathscr{D}_t^* \quad \text{and} \quad S(t)^*x' = S(t^*)x',$$

i.e.

$$\text{Rstr.}_{\mathscr{D}} S(t^*) \subseteq S(t)^*.$$

Hence to complete the proof of (α) we need only prove (II).

Proof of (II). Fix $x' \in \mathscr{D}$, say $x' = X(\lambda')(w')$. Now for $x \in \mathscr{D}$, say $x = X(\lambda)(w)$, we find that

$$\begin{aligned}(S(t)x, x') &= (S(t)X(\lambda)w, X(\lambda')w') \\ &= [(X(\lambda')^* \cdot X(t \oplus \lambda)(w))](w') && \text{cf. 2.2(c) and (6)} \\ &= [(X(t^* \oplus \lambda')^* \cdot X(t)(w))](w') && \text{by (3)} \\ &= (X(t)w, X(t^* \oplus \lambda')w') && \text{cf. 2.2(c)} \\ &= (X(t)w, S(t^*)X(\lambda')w') && \text{by (6)} \\ &= (x, S(t^*)x').\end{aligned}$$

This shows that (II) holds $\forall x \in \mathscr{D}$, whence it follows easily that (II) holds $\forall x \in \langle \mathscr{D} \rangle$. Now let $x \in \mathscr{D}_t$. By (1′) $\exists$ a sequence $(x_n)_1^\infty$ in $\langle \mathscr{D} \rangle$ such that $x_n \to x$ and $S(t)x_n \to S(t)x$. Replacing x by x_n in (II), we conclude at once on letting $n \to \infty$ that (II) holds for x. Thus (II) is established.

This completes the proof of (α). □

The following corollary of Thm. 4.7 plays an ancillary role in our quest for continuous propagators.

4.8 COROLLARY. *Let* $\Lambda, W, \mathfrak{H}, \mathscr{D}, \Gamma$ *be as in* (4.6), *and let* (4.5)(a) *prevail. Then for the* $S(\cdot)$ *obtained in Thm.* 4.7 *we have*

(a) $\forall t \in \Gamma$ *and* $\forall x, x' \in \langle \mathscr{D} \rangle$, $(S(t)x, x') = (x, S(t^*)x')$;

(b) $(\text{Rstr.}_{\langle \mathscr{D} \rangle} S(t): t \in \Gamma$ *is a semi-group of linear operators on* $\langle \mathscr{D} \rangle$ *to* $\langle \mathscr{D} \rangle$ *and*

$$S(t^*) = S(t)^* \quad \textit{on } \langle \mathscr{D} \rangle;$$

(c) $\forall x \in \langle \mathscr{D} \rangle$ *and* $\forall t \in \Gamma$,

$$|S(t)x|^2 \leqq |S(t^*)S(t)x|\,|x| = |S(t^* + t)x|\,|x|;$$

(d) $\forall x \in \langle \mathscr{D} \rangle$, $\forall p \in \Gamma_{0+}$ *and* $\forall n \in \mathbb{N}_+$,

$$|S(p)x|^{2n} \leqq |S(2^n p)x| \cdot |x|^{2n-1}.$$

Proof. (a) Let $t \in \Gamma$ and $x, x' \in \langle \mathscr{D} \rangle$. As (4.5)(a) holds, we know from 4.7 that

$$\text{Rstr.}_{\langle \mathscr{D} \rangle} S(t^*) \subseteq S(t)^*.$$

This means that the equality in (a) holds $\forall x \in \langle \mathscr{D} \rangle$ and $\forall x' \in \mathscr{D}$. From this it follows easily that the equality also holds for all x' which are linear combinations of elements of $\mathscr{D}$, i.e. $\forall x' \in \langle \mathscr{D} \rangle$.

(b) Define S_t, $\langle \mathscr{D} \rangle$ as in the proof of 4.7. Then

$$\forall t \in \Gamma \text{ and } \forall \lambda \in \Lambda, \qquad S_t \cdot X(\lambda) = X(t \oplus \lambda). \tag{1}$$

Hence

$$S_{s+t} \cdot X(\lambda) = X\{(s+t) \oplus \lambda\} = X\{s \oplus (t \oplus \lambda)\} = S_s \cdot X(t \oplus \lambda) = (S_s \cdot S_t) \cdot X(\lambda).$$

The domain of each $S_\tau \circ X(\lambda)$ being W, it follows that

$$S_{s+t} = S_s \circ S_t \quad \text{on} \bigcup_{\lambda \in \Lambda} X(\lambda)(W), \quad \text{i.e. on } \mathscr{D}.$$

From this it follows in turn, since Range $S_t \subseteq \mathscr{D} =$ domain of S_s, and $\langle S_s \rangle$ is single-valued, that

$$\langle S_{s+t} \rangle = \langle S_s \rangle \cdot \langle S_t \rangle \subseteq \mathscr{S}_X \times \mathscr{S}_X,$$

i.e. that

$$S(s+t) = S(s) \circ S(t) \quad \text{on } \langle \mathscr{D} \rangle.$$

Taking $t = 0$ in (1), we obtain $S_0 = I_{\mathscr{D}}$, whence $\langle S_0 \rangle = I_{\langle \mathscr{D} \rangle}$, i.e. $S(0) = I_{\mathscr{S}_X}$ on $\langle \mathscr{D} \rangle$. Finally, (a) tells us that $S(t^*) = S(t)^*$ on $\langle \mathscr{D} \rangle$. Thus (b) is proved.

(c) Let $x \in \langle \mathscr{D} \rangle$ and $t \in \Gamma$. Then $x' =_d S(t)x \in \langle \mathscr{D} \rangle$, and so by (a), the Schwarz inequality and (b),

$$|S(t)x|^2 = (S(t)x, x') = (x, S(t^*)x') \leqq |x| \, |S(t^*)S(t)x| = |x| \, |S(t^* + t)x|.$$

(d) Now let $p \in \Gamma_{0+}$ and $x \in \langle \mathscr{D} \rangle$. Then by (c) and the fact that $p = p^*$, cf. (4.3), we get

$$|S(p)x|^2 \leqq |x| \cdot |S(2p)x|. \tag{2}$$

Thus (d) holds for $n = 1$. By (2),

$$|S(p)x|^{2^{n+1}} = (|S(p)x|^2)^{2^n} \leqq |x|^{2^n} \cdot |S(2p)x|^{2^n}. \tag{3}$$

Now grant the result (d) for n, and replace p by $2p$ in it; this yields

$$|S(2p)x|^{2^n} \leqq |x|^{2^n - 1} |S(2^n 2p)x|.$$

Substitution on the RHS(3) yields the result (d) for $n+1$. By induction (d) holds for all $n \in \mathbb{N}_+$. □

We now impose‡ both the requirements (4.5)(a), (b) on $K(\cdot \cdot)$, and assert the continuity of the operator $S(t)$ restricted to the (nonlinear) set $\mathscr{D}$, as well as some equalities governing $|S(t)|$:

4.9 Lemma. *Let* Λ, W, $\mathfrak{H}$, $X(\cdot)$, $\mathscr{D}$, Γ *be as in* (4.6), *and let both* (4.5)(a), (b) *prevail. Then for the* $S(\cdot)$ *obtained in Thm.* 4.7, *we have*

(a) $\forall t \in \Gamma$,

$$\alpha(t) =_d \sup_{0 \neq x \in \mathscr{D}} (|S(t)x| / |x|) \leqq \sqrt{\gamma(t)} < \infty;$$

(b) $\forall t \in \Gamma$,

$$\alpha(t^*) = \alpha(t), \qquad \alpha(t^* + t) = \alpha(t)^2;$$

‡ See the first part of Note 2 added in proof.

(c) $\forall p \in \Gamma_{0+}$ *and* $\forall n \in \mathbb{N}_+$,

$$\alpha(p)^{2^n} = \alpha(2^n p).$$

Proof. (a) Fix $t \in \Gamma$ and consider any $x \in \mathcal{D}$, say $x \underset{d}{=} X(\lambda)(w)$. Then by 2.10(b) and (4.5)(b)

$$|S(t)x| = |S(t)X(\lambda)(w)| = |X(t\oplus\lambda)w| \leqq \sqrt{\{K(t\oplus\lambda, t\oplus\lambda)(w)(w)\}}$$

$$\leqq \sqrt{\{\gamma(t)K(\lambda,\lambda)(w)(w)\}} = \sqrt{\gamma(t)} \cdot |x|.$$

It follows that $\alpha(t) \leqq \sqrt{\gamma(t)} < \infty$.

(b) Again fix $t \in \Gamma$, and consider any $x \in \mathcal{D}$. By 4.8(c)

$$(1) \qquad |S(t)x|^2 \leqq |S(t^*)S(t)x|\,|x| \leqq \alpha(t^*)|S(t)x|\,|x|.$$

On canceling $|S(t)x|$ we easily get $\alpha(t) \leqq \alpha(t^*)$. The replacement of t by t^* yields $\alpha(t^*) \leqq \alpha(t)$. Thus

$$(2) \qquad \alpha(t^*) = \alpha(t).$$

Again from 4.8(c)

$$|S(t)x|^2 \leqq |S(t^*+t)x| \cdot |x| \leqq \alpha(t^*+t)|x|^2,$$

whence $\alpha(t)^2 \leqq \alpha(t^*+t)$. But from the s.g. property, (1) and (2),

$$|S(t^*+t)x| = |S(t^*)S(t)x| \leqq \alpha(t^*)\alpha(t)|x| = \alpha(t)^2|x|,$$

whence $\alpha(t^*+t) \leqq \alpha(t)^2$. Thus (b) is proved.

(c) Let $p \in \Gamma_{0+}$ and $n \in \mathbb{N}_+$. Then 4.8(d) at once yields

$$\alpha(p)^{2^n} \leqq \alpha(2^n p).$$

On the other hand, from $S(nt)x = S(t)^n x$, $x \in \mathcal{D}$, we get

$$|S(nt)x| \leqq \alpha(t)|S(t)^{n-1}x| \leqq \cdots \leqq \alpha(t)^n |x|,$$

whence $\alpha(nt) \leqq \alpha(t)^n$. In particular,

$$\alpha(2^n p) \leqq \alpha(p)^{2^n}.$$

Thus (c) is proved. □

4.10 Main Theorem (existence of propagator). *Let* (i) X *be a function on* Λ *to* $\mathrm{CL}(W, \mathfrak{H})$, *where* Λ *is any set,* W *is a Banach space over* $\mathbb{F}$ *and* $\mathfrak{H}$ *is a Hilbert space over* $\mathbb{F}$, (ii) Γ *be a (not necessarily Abelian) involutory semi-group with neutral element* 0 *that acts on* Λ (cf. Def. 3.2(a)). *Then*

(a) *the following conditions are equivalent:*

(α) $X(\cdot)$ *possesses a propagator* $S(\cdot)$ *on* Γ *to* $\mathrm{CL}(\mathcal{S}_X, \mathcal{S}_X)$ *such that*

$$\forall t \in \Gamma, \qquad S(t^*) = S(t)^*,$$

(β) *the covariance kernel* $K(\cdot\,\cdot)$ *of* $X(\cdot)$ *possesses the transfer and mild translational properties* (4.5)(a), (b);

(b) *when* (β) *holds and* $\gamma(\cdot)$ *is as in* (4.5)(b), *we have* $\forall t \in \Gamma$, $|S(t)| \leqq \sqrt{\gamma(t)}$; *when* (α) *holds, the best possible choice for* $\gamma(\cdot)$ *in* (4.5)(b) *is* $\gamma(t) = |S(t)|^2$.

Proof. (a) Let (α) hold, and $\lambda, \lambda' \in \Lambda$ and $t \in \Gamma$. Then obviously

$$K(\lambda, t^*\oplus\lambda') = X(t^*\oplus\lambda')^* \cdot X(\lambda) = X(\lambda')^* \cdot S(t^*) \cdot X(\lambda)$$

$$= X(\lambda')^* \cdot S(t)X(\lambda) = X(\lambda')^* \cdot X(t\oplus\lambda) = K(\lambda', t\oplus\lambda).$$

i.e. we have (4.5)(a). Also,

$$K(t\oplus\lambda, t\oplus\lambda) = X(t\oplus\lambda)^* \cdot X(t\oplus\lambda) = X(\lambda)^* \cdot S(t)^* \cdot S(t)X(\lambda).$$

Now $S(t)^* \cdot S(t)$ being nonnegative hermitian on $\mathscr{S}_X$ to $\mathscr{S}_X$ with norm $|S(t)|^2$, we have

$$0 \preceq S(t)^* \cdot S(t) \preceq |S(t)|^2 I_{\mathscr{S}_X}.$$

Hence, cf. 2.4(b),

$$0 \preceq X(\lambda)^* S(t)^* S(t) X(\lambda) \preceq |S(t)|^2 K(\lambda, \lambda).$$

Thus

$$K(t\oplus\lambda, t\oplus\lambda) \preceq |S(t)|^2 K(\lambda, \lambda),$$

i.e. we have (4.5)(b) with $\gamma(t) = |S(t)|^2$. Thus (β) holds.

Next let (β) hold. Then by (4.5)(a) and Thm. 4.7, $\forall t \in \Gamma$, $\exists$ a (s.v.) closed linear operator $S(t)$ from $\mathscr{S}_X$ to $\mathscr{S}_X$ with domain $\mathscr{D}_t \supseteq \langle\mathscr{D}\rangle$, i.e. with an e.d. domain in $\mathscr{S}_X$, such that

$$\text{(1)} \qquad \forall t \in \Gamma \text{ and } \forall \lambda \in \Lambda, \qquad S(t) \circ X(\lambda) = X(t \oplus \lambda).$$

But now that (4.5)(b) prevails, we also claim that

$$\text{(I)} \qquad \forall t \in \Gamma, \qquad \beta(t) \underset{d}{=} \sup_{0 \neq x \in \langle\mathscr{D}\rangle} \frac{|S(t)x|}{|x|} = \sup_{0 = x \in \mathscr{D}} \frac{|S(t)x|}{|x|} \underset{d}{=} \alpha(t).$$

Proof of (I). Fix $t \in \Gamma$, and consider any $x \in \langle\mathscr{D}\rangle$. Since $\mathscr{D}$ is closed under multiplication by scalars, we have $x = \sum_1^r x_i$, where $x_i \in \mathscr{D}$, and so by 4.9(a)

$$\text{(2)} \qquad |S(t)x| = \left|\sum_1^r S(t)x_i\right| \leqq \sum_1^r |S(t)x_i| \leqq \alpha(t) \sum_1^r |x_i| = \alpha(t) c_x,$$

where $c_x > 0$ is independent of t.

Now suppose $\beta(t) \neq \alpha(t)$. Since $\mathscr{D} \subseteq \langle\mathscr{D}\rangle$ and therefore $\alpha(t) \leqq \beta(t)$, it follows that $\alpha(t) < \beta(t) \leqq \infty$. Hence

$$\exists r > 1, \text{ and } \exists x \in \langle\mathscr{D}\rangle \ni \quad |x| = 1 \quad \text{and} \quad r\alpha(t) < |S(t)x|.$$

Letting $p \underset{d}{=} t^* + t$, it follows from 4.8(c) that

$$r^2 \alpha(t)^2 < |S(t)x|^2 \leqq |S(t^* + t)x| = |S(p)x|,$$

whence from 4.8(d) and (2)

$$\text{(3)} \qquad \{r^2 \alpha(t)^2\}^{2^n} \leqq |S(p)x|^{2^n} \leqq |S(2^n p)x| \leqq \alpha(2^n p) \cdot c_x.$$

But by 4.9(c) and (b),

$$\alpha(2^n p) = \alpha(p)^{2^n} = \alpha(t^* + t)^{2^n} = \{\alpha(t)^2\}^{2^n} = \alpha(t)^{2^{n+1}}.$$

Thus (3) reduces to

$$r^{2^{n+1}} \alpha(t)^{2^{n+1}} \leqq \alpha(t)^{2^{n+1}} \cdot c_x, \quad \text{i.e.} \quad r^{2^{n+1}} \leqq c_x.$$

But this is impossible since $r > 1$ and $r^{2^{n+1}} \to \infty$, as $n \to \infty$. Hence (I) is proved.

Condition (I) shows of course that

$$\text{Rstr.}_{\langle\mathscr{D}\rangle} S(t) \in \text{CL}(\langle\mathscr{D}\rangle, \langle\mathscr{D}\rangle).$$

From the closedness of $S(t)$ and the extension principle for continuous linear operators it follows, since $\mathcal{S}_X = \text{cls.}\langle\mathcal{D}\rangle$, that

(4) $$S(t) \supseteq \text{cls. Rstr.}_{\langle\mathcal{D}\rangle} S(t) \in \text{CL}(\mathcal{S}_X, \mathcal{S}_X) \quad \text{i.e.} \quad S(t) \in \text{CL}(\mathcal{S}_X, \mathcal{S}_X).$$

By (4) and (1), $S(\cdot)$ is a propagator of $X(\cdot)$.

Finally, since by (4) $S(t)$ and $S(t^*)$ are continuous on $\mathcal{S}_X$, the equality in 4.8(b) valid on $\langle\mathcal{D}\rangle$ holds throughout cls. $\langle\mathcal{D}\rangle = \mathcal{S}_X$; i.e.

$$\forall x, x' \in \mathcal{S}_X, \qquad (S(t)x, x') = (x, S(t^*)x').$$

This shows that $S(t^*) = S(t)^*$, and the proof of (α) is over.

(b) Let (β) hold. Then from the extension principle for continuous linear operators, and (I) and 4.9(a), we have

$$|S(t)| = \beta(t) = \alpha(t) \leqq \sqrt{\gamma(t)}.$$

Next let (α) hold. Then as just shown in the derivation of (β), we can take $\gamma(t) = |S(t)|^2$. This is the best choice of $\gamma(t)$ in view of the last inequality. Thus we have (b). □

The following result states how the nature of the *s.g. Γ affects that of the propagator $S(\cdot)$. It is an obvious consequence of the assertion $S(t^*) = S(t)^*$ in the last theorem.

4.11 COROLLARY. *Let* (i) $\Lambda, W, \mathfrak{H}, X(\cdot), \Gamma$ *be as in* (4.6); (ii) *the covariance kernel* $K(\cdot\,\cdot)$ *of* $X(\cdot)$ *have the transfer and translational properties* (4.5)(a), (b); (iii) $S(\cdot)$ *be the propagator of* $X(\cdot)$ *on* Γ *to* $\text{CL}(\mathcal{S}_X, \mathcal{S}_X)$. *Then* $\forall t \in \Gamma$, *we have* (cf. (4.2))

(a) $$t \textit{ is normal} \Rightarrow S(t) \textit{ is normal},$$
$$t \textit{ is s.a.} \Rightarrow S(t) \textit{ is s.a.},$$
$$t \geq 0 \Rightarrow S(t) \geq 0;$$

(b) $$t^* = t = 2t \Rightarrow S(t) \textit{ is a } (\perp) \textit{ projection};$$

(c) $$t^* + t = 0 \Rightarrow S(t) \textit{ is isometric},$$
$$t^* + t = 0 = t + t^* \Rightarrow S(t) \textit{ is unitary};$$

(d) Γ *is a group and* $t^* = -t \Rightarrow S(\cdot)$ *is a group of unitary operators on* $\mathcal{S}_X$ *onto* $\mathcal{S}_X$.

These relations correspond to the (i)–(iv) of Nagy [26, p. 21] for whom $\Lambda = \Gamma$, W is a Hilbert space, and the covariance $K(\cdot\,\cdot)$ of $X(\cdot)$ is derived from a PD function $R(\cdot)$ as in 4.1(b). In view of the frequent occurrence of the case $\Lambda = \Gamma$, it is convenient to introduce the following concept and to rephase Thm. 4.10 in terms of it.

4.12 DEFINITION. Let (i) Λ be an (additive) *s.g. with neutral element 0, (ii) W be a Banach space over $\mathbb{F}$ and $\mathfrak{H}$ a Hilbert space over $\mathbb{F}$, (iii) $X(\cdot)$ be a function on Λ to $\text{CL}(W, \mathfrak{H})$. We say that $X(\cdot)$ has the *covariance function* $R(\cdot)$ on Λ to $\text{Cl}(W, W^*)$ iff

$$\forall \lambda, \lambda' \in \Lambda, \qquad R(\lambda'^* + \lambda) = X(\lambda')^* X(\lambda).$$

4.13 MAIN THEOREM (existence of propagator).§ *Let* $\Lambda, W, \mathfrak{H}, X(\cdot)$ *be as in* 4.12. *Then*

(a) *the following conditions are equivalent*:

(α) $X(\cdot)$ *possesses a propagator* $S(\cdot)$ *on* Λ *to* $\text{CL}(\mathcal{S}_X, \mathcal{S}_X)$ *such that*

$$\forall t \in \Lambda, \qquad S(t^*) = S(t)^*,$$

§ See the second part of Note 2 added in proof.

(β) *$X(\cdot)$ possesses a covariance function $R(\cdot)$ having the translational property*: ∃ *a function $\gamma(\cdot)$ on Λ to $\mathbb{R}_{0+}$* ∋

$$\forall \lambda, t \in \Lambda, \qquad 0 \leq R(\lambda^* + t^* + t + \lambda) \leq \gamma(t) \cdot R(\lambda^* + \lambda);$$

(b) *when* (β) *prevails, we have* $\forall t \in \Lambda$, $|S(t)| \leqq \sqrt{\gamma(t)}$; *when* (α) *holds, the best choice for the* $\gamma(\cdot)$ *in* (β) *is* $\gamma(t) = |S(t)|^2$.

Proof. The possession by $X(\cdot)$ of a covariance function is equivalent to its possession of a covariance kernel $K(\cdot\,\cdot)$ having the transfer property (4.5)(a); and the inequality in (β) is equivalent to that in (4.5)(b). Hence 4.13 is a mere rephrasing of Thm. 4.10 in the setting $\Lambda = \Gamma$. □

Next, we consider the case in which the parameter space Λ is a Banach*-algebra $\mathbb{A}$ over $\mathbb{F}$ with unit **1**. $\mathbb{A}$ is of course a *s.g. under multiplication with neutral element **1**. A PD function $R(\cdot)$ on (this *s.g.) $\mathbb{A}$ to CL(W, W^*) would come under the scope of Thm. 4.13 (with + in 4.13 interpreted as multiplication) were $R(\cdot)$ to satisfy the inequality in 4.13(β). We shall now show that if the PD function $R(\cdot)$ on $\mathbb{A}$ (considered as a vector space) is *linear*, $|\mathbf{1}| = 1$ and $|a^*| = |a|$, then the inequality does indeed prevail. Furthermore, from the kernel theorem in its vectorial format 2.16 we shall deduce that the variety $X(\cdot)$ with covariance function $R(\cdot)$ is itself a continuous linear operator on (the Banach space) $\mathbb{A}$, and thence by appeal to Thm. 4.13 that so is its propagator $S(\cdot)$. Our theorem reads as follows:

4.14 THEOREM. *Let* (i) $\mathbb{A}$ *be a Banach algebra over* $\mathbb{F}$ *with unit* $\mathbf{1} \ni |\mathbf{1}| = 1$ *and an isometric involution* *; (ii) $R(\cdot)$ *be a linear operator on* (*the vector space*) $\mathbb{A}$ *to* CL(W, W^*), *where W is a Banach space over* $\mathbb{F}$; (iii) $R(\cdot)$ *be a* PD *function on* (*the multiplicative* *s.g.) $\mathbb{A}$ *to* CL(W, W^*); (iv) $X(\cdot)$ *be the variety on* $\mathbb{A}$ *to* CL(W, $\mathfrak{H}$) *with covariance function* $R(\cdot)$, $\mathfrak{H}$ *being as in the Kernel Thm.* 2.10. *Then*

(a) $R(\cdot)$ *is continuous on* $\mathbb{A}$, $|R| = |R(\mathbf{1})|$, *and*

$$\forall a, t \in \mathbb{A}, \qquad 0 \leq R(a^* t^* t a) \leq |t|^2 R(a^* a);$$

(b) $X(\cdot) \in \mathrm{CL}(\mathbb{A}, \mathrm{CL}(W, \mathfrak{H}))$ *and* $|X| = \sqrt{|R(\mathbf{1})|} = |X(\mathbf{1})|$;

(c) $X(\cdot)$ *possesses a propagator* $S(\cdot)$ *on* $\mathbb{A}$ *to* CL($\mathcal{S}_X$, $\mathcal{S}_X$), *i.e.*

$$\forall a, t \in \mathbb{A}, \qquad X(ta) = S(t) \cdot X(a);$$

moreover, $\forall t \in \mathbb{A}$, $S(t^*) = S(t)^*$, *and* $S(\cdot)$ *is a linear contraction of norm* 1 *on* $\mathbb{A}$ *to* CL($\mathcal{S}_X$, $\mathcal{S}_X$); *thus the propagator* $S(\cdot)$ *is a* *-*representation of* $\mathbb{A}$ *in* CL($\mathcal{S}_X$, $\mathcal{S}_X$).[16]

Proof. (a) Write "$a \preceq b$" for "$b - a \in \mathbb{A}_{0+}$". Then, cf. [28, pp. 188–189], $\forall p \in \mathbb{A}_{0+}$ and $\forall \rho \geqq |p|$, $0 \preceq p \preceq \rho \mathbf{1}$, whence $\forall a \in \mathbb{A}$, $0 \preceq a^* p a \preceq \rho a^* a$. Taking $p = t^* t$, $\rho = |t|^2$ and noting that $|t|^2 = |t^*||t| \geqq |t^* t|$, we get

(1)
$$\forall a, t \in \mathbb{A}, \qquad 0 \preceq a^* t^* t a \preceq |t|^2 a^* a.$$

Now (iii) tells us that the kernel $K(\cdot\,\cdot)$ given by

(2)
$$\forall a, b \in \mathbb{A}, \qquad K(a, b) \underset{d}{=} R(b^* a)$$

is PD on $\mathbb{A} \times \mathbb{A}$ to CL(W, W^*). From 2.7(a), (c) we therefore conclude that $\forall a, t \in \mathbb{A}$,

(3)
$$0 \preceq R(t^* t) \quad \text{and} \quad |R(a^* t)|^2 \leqq |R(a^* a)| \cdot |R(t^* t)|.$$

Thus the linear operator $R(\cdot)$ carries nonnegative elements of $\mathbb{A}$ into nonnegative elements of CL(W, W^*), and is therefore "nondecreasing" with respect to the relation

[16] That is, $S(\cdot)$ is a continuous involution-preserving multiplicative linear operator on $\mathbb{A}$ to CL($\mathcal{S}_X$, $\mathcal{S}_X$). It easily follows that $S(\cdot)$ is a cyclic representation when dim $W = 1$, e.g. when $W = \mathbb{F}$.

$\preceq$. Hence from (1), we get

(4) $$\forall a, t \in \mathbb{A}, \qquad 0 \preceq R(a^*t^*ta) \preceq |t|^2 R(a^*a).$$

Taking $a = 1$ and applying 2.4(g), we get $|R(t^*t)| \leqq |t|^2 \cdot |R(\mathbf{1})|$. So, from (3) with $a = 1$, it follows that

$$|R(t)|^2 \leqq |R(\mathbf{1})|\,|R(t^*t)| \leqq |t|^2 |R(\mathbf{1})|^2,$$

whence $|R(t)| \leqq |t|\,|R(\mathbf{1})|$, and so

$$|R| \leqq |R(\mathbf{1})| = \frac{|R(\mathbf{1})|}{|\mathbf{1}|} \leqq |R|, \quad \text{i.e.,} \quad \|R| = \|R(\mathbf{1})|.$$

We thus have (a).

(b) From the distributive laws of $\mathbb{A}$ and the linearity of $R(\cdot)$ on $\mathbb{A}$, we easily see that the kernel $K(\cdot\,\cdot)$ in (2) is sesquilinear on $\mathbb{A} \times \mathbb{A}$ to CL(W, W^*), and taking $a = \mathbf{1}$ in (a) and using 2.4(g), we get

$$0 \leqq |K(t, t)| \leqq |t|^2 |R(\mathbf{1})|.$$

Hence $K(\cdot\,\cdot)$ is Lipschitzian, and since the last inequality becomes an equality for $t = \mathbf{1}$, we have $|K| = |R(\mathbf{1})|$. Now by (iv) and 4.1(a), $K(\cdot\,\cdot)$ is the covariance kernel of $X(\cdot)$. Hence by 2.16(b), $|X| = \sqrt{|K|} = \sqrt{|R(\mathbf{1})|}$. Thus (b) holds.

(c) From (a) we see that $R(\cdot)$ satisfies the translational inequality in 4.13(β) with $\gamma(t) = |t|^2$. Hence by 4.13, $X(\cdot)$ possesses a propagator $S(\cdot)$ on $\mathbb{A}$ to CL$(\mathscr{S}_X, \mathscr{S}_X)$ such that

(5) $$\forall t \in \mathbb{A}, \quad S(t^*) = S(t)^* \quad \text{and} \quad |S(t)| \leqq |t|.$$

We now assert

(I) $$\forall a, b \in \mathbb{A}, \quad \forall \alpha, \beta \in \mathbb{F}, \quad S(\alpha a + \beta b) = \alpha S(a) + \beta S(b).$$

Proof of (I). Let $\lambda \in \mathbb{A}$. Then by fully exploiting the propagation equality for $S(\cdot)$ and the linearity of $X(\cdot)$ we get

$$S(\alpha a + \beta b) \cdot X(\lambda) = \{\alpha S(a) + \beta S(b)\} \cdot X(\lambda).$$

Thus for the operator

$$D \underset{d}{=} S(\alpha a + b\beta) - \alpha S(a) - \beta S(b)$$

we find that $D \cdot X(\lambda) = 0$. As this holds $\forall \lambda \in \Lambda$, we conclude that

$$\mathscr{D} \underset{d}{=} \bigcup_{\lambda \in \mathbb{A}} X(\lambda)(W) \subseteq \mathscr{N} \underset{d}{=} \text{null space of } D.$$

Since D is linear and continuous on $\mathscr{S}_X$, it follows that $\mathscr{S}_X \underset{d}{=} \text{cls.}\,\langle \mathscr{D} \rangle \subseteq \mathscr{N}$, i.e. $D = 0$. This proves (I).

Finally, $S(\cdot)$ being a semi-group, cf. 3.3, we have

(6) $$\forall a, b \in \mathbb{A}, \quad S(ab) = S(a) \cdot S(b), \quad S(1) = I_{\mathscr{S}_X} \quad \text{and} \quad |S(\mathbf{1})| = 1 = |\mathbf{1}|.$$

By (5), (I), (6) we have (c). □

The last theorem is restateable in the following form in which it can be directly compared to an important theorem which Stinespring proved ab initio in 1955 [32, Thm. 1]:

4.15 THEOREM. *Let* (i) *and* (ii) *be as in Thm.* 4.14. *Then*

(a) *the necessary and sufficient condition that*

$$\forall a \in \mathbb{A}, \qquad R(a) = J^* S(a) J,$$

where $J \in \mathrm{CL}(W, \mathfrak{H}_0)$, $\mathfrak{H}_0$ *being a Hilbert space over* $\mathbb{F}$, *and* $S(\cdot)$ *is a* **-representation of* $\mathbb{A}$ *into* $\mathrm{CL}(\mathfrak{H}_0, \mathfrak{H}_0)$ *is that* $R(\cdot)$ *be* PD *on* (*the multiplicative* *s.g.) $\mathbb{A}$;

(b) *when* $R(\cdot)$ *is* PD *on* $\mathbb{A}$, *and* $X(\cdot)$ *is a variety on* $\mathbb{A}$ *with covariance function* $R(\cdot)$ (cf. 4.12 and 2.10) *then we can in* (a) *take* $\mathfrak{H}_0 = \mathcal{S}_X$, $S(\cdot) =$ *the propagator of* $X(\cdot)$ *and* $J = X(\mathbf{1})$.

Proof. (a) *Sufficiency*. Let $R(\cdot)$ be PD on $\mathbb{A}$, and $X(\cdot)$ be as described in (b). Then all the premises (i)–(iv) of Thm. 4.14 are fulfilled. Hence by Thm. 4.14, $X(\cdot)$ has a propagator $S(\cdot)$ which is a *-representation of $\mathbb{A}$ in $\mathrm{CL}(\mathcal{S}_X, \mathcal{S}_X)$. With $\mathfrak{H}_0$ and J as described in (b), the *-representation is in $\mathrm{CL}(\mathfrak{H}_0, \mathfrak{H}_0)$, and

$$J^*S(a)J = X(\mathbf{1})^*S(a)X(\mathbf{1}) = X(1)^* \cdot X(a) = R(a),$$

as desired.

Necessity. Let J, $\mathfrak{H}_0$, $S(\cdot)$ be as described in (a) and define $\forall a \in \mathbb{A}$, $Y(a) \underset{d}{=} S(a)J$. Then $Y(\cdot)$ is a variety on $\mathbb{A}$ to $\mathrm{CL}(W, \mathfrak{H}_0)$. Since $S(\cdot)$ is a *-representation of $\mathbb{A}$, we have $\forall a, b \in \mathbb{A}$,

$$Y(b)^* Y(a) = J^* \cdot S(b)^* \cdot S(a)J = J^* \cdot S(b^*a)J = R(b^*a).$$

Thus $R(\cdot)$ is the covariance function of $Y(\cdot)$ and therefore obviously $R(\cdot)$ is a PD function.

(b) This has been shown in the sufficiency part of the proof of (a). □

4.16 *Remarks*. The only difference between Thm. 4.15 and Stinespring's [32, Thm. 1] consists in our saying "Banach algebra with unit of norm 1 having isometric involution" instead of "C^*-algebra"; "Banach space" instead of "Hilbert space"; and "PD on $\mathbb{A}$" instead of "completely positive on $\mathbb{A}$". Now in his proof of Thm. 1, Stinespring shows that *a linear operator* $R(\cdot)$ *on* $\mathbb{A}$ *to* $\mathrm{CL}(W, W)$, ($W =$ *Hilbert space*) *is* PD *on* $\mathbb{A}$ *iff* $R(\cdot)$ *is "completely positive"* (Def. in [32, p. 211]). If we treat this result as a separate lemma, then we may regard 4.15 as an explication and generalization of Stinespring's Thm. 1. Thm. 4.15 is more explicit in that in (b) the auxiliary Hilbert space $\mathfrak{H}_0$, the *-representation $S(\cdot)$ and the operator J are identified. It is a generalization in that it allows W to be a Banach (rather than Hilbert) space, and $\mathbb{A}$ to be isometric-involutory (rather than C^*).

Another important theorem of Stinespring [32, Thm. 4] asserts that *for a commutative* C^**-algebra* $\mathbb{A}$ *every positive linear operator* R *on* $\mathbb{A}$ *to* $\mathrm{CL}(W, W)$, $W =$ *a Hilbert space, is a* PD *function on the* * s.g. R is "*positive*" means of course that $a \geq 0$ in $\mathbb{A} \Rightarrow R(a) \geq 0$ in $\mathrm{CL}(W, W)$. In the light of this we can obviously assert the following:

4.17 COROLLARY. *Let* (i) $\mathbb{A}$ *be a commutative* C^**-algebra with unit* $\mathbf{1}$, (ii) $R(\cdot)$ *be a positive linear operator on* $\mathbb{A}$ *to* $\mathrm{CL}(W, W)$, *where* W *is a Hilbert space. Then the conclusions* (a)–(c) *of Thm.* 4.14 *are valid*; *moreover*, $\forall t \in \mathbb{A}$, $S(t)$ *is normal.*

5. Dilation theorems. So far W has been a Banach space, and the parameter space Λ any arbitrary set, considered either singly or along with a semi-group Γ acting on it. For dilation theory (as now conceived) W has to be a Hilbert space, and Λ itself has to be a semi-group. In this setting we have the following general result which shows that all propagators are dilations and vice versa:

5.1 GENERAL DILATION THEOREM. *Let*

(i) Λ *be a semi-group under* $+$ (*not necessarily Abelian*) *with neutral element* 0,

(ii) W *be a Hilbert space over* $\mathbb{F}$,

(iii) $K(\cdot\,\cdot)$ *be a* PD *kernel on* $\Lambda \times \Lambda$ *to* $\mathrm{CL}(W, W)$ *for which* $K(0, 0) = I_W$.

Then

(a) *if a variety* $X(\cdot)$ *on* Λ *to* $\mathrm{CL}(W, \mathfrak{H})$ *the covariance kernel of which is* $K(\cdot\,\cdot)$, $\mathfrak{H}$ *being as in* 2.10, *has the propagator* $S(\cdot)$, *then* $S(-)^* \cdot S(\cdot)$ *is the dilation of* $K(\cdot, -)$ *in the Hilbert space* $\mathcal{S}_X \subseteq \mathfrak{H}$;

(b) *if* $S(\cdot)$ *is a semi-group on* Λ *to* $\mathrm{CL}(\mathfrak{H}, \mathfrak{H})$ *and* $S(-)^* \cdot S(\cdot)$ *is a dilation of* $K(\cdot, -)$, *then a restriction* $S_0(\cdot) \subseteq S(\cdot)$ *is the propagator of a variety* $X(\cdot)$ *having* $K(\cdot\cdot)$ *as its covariance kernel.*

Proof. (a) If $S(\cdot)$ is the propagator of $X(\cdot)$, then by Def. 3.2(b), $\forall \lambda \in \Lambda$, $X(\lambda) = S(\lambda)X(0)$. Hence

$$K(\lambda, \lambda') = X(\lambda')^* \cdot X(\lambda) = X(0)^* \cdot S(\lambda')^* \cdot S(\lambda) \cdot X(0).$$

Since $I_W = K(0, 0) = X(0)^* \cdot X(0)$, i.e. $X(0)$ is an isometry on W to $\mathscr{S}_X$, it follows by Def. 1.2 that $S(-)^*S(\cdot)$ is a dilation of $K(\cdot, -)$.

(b) By Def. 1.2, $\exists$ an isometry J on W to a Hilbert space $\mathfrak{H}$ such that

$$(1) \qquad \forall \lambda, \lambda' \in \Lambda, \qquad K(\lambda, \lambda') = J^* S(\lambda')^* S(\lambda) J.$$

Clearly the variety $X(\cdot) \underset{d}{=} S(\cdot)J$ on Λ to $\mathrm{CL}(W, \mathfrak{H})$ has $K(\cdot, -)$ as covariance kernel. And, since $S(\cdot)$ is a semi-group,

$$\forall t, \lambda \in \Lambda, \qquad X(t+\lambda) \underset{d}{=} S(t+\lambda)J = S(t)S(\lambda)J = S(t)X(\lambda)$$

and

$$S(\lambda)(\mathscr{S}_X) \subseteq \mathscr{S}_X.$$

Hence, cf. Def. 3.2(b), $S_0(\cdot) \underset{d}{=} \mathrm{Rstr.}_{\mathscr{S}_X} S(\cdot)$ is a propagator of $X(\cdot)$. □

The next result tells us when the dilation will be strongly continuous.

5.2 COROLLARY (strongly continuous dilation). *Let*

(i) Λ, W, $K(\cdot\cdot)$, $X(\cdot)$ *and* $S(\cdot)$ *be as in* 5.1,

(ii) *the semi-group* Λ *in* (i) *be topological* (cf. 3.6(ii)),

(iii) *the* PD *kernel* $K(\cdot\cdot)$ *be strongly continuous on* $\Lambda \times \Lambda$, *i.e.*

$$\forall \lambda, \lambda' \in \Lambda, \quad \operatorname*{slim}_{(t,t') \to (0,0)} K(t+\lambda, t'+\lambda') = K(\lambda, \lambda'),$$

(iv) *the function* $\gamma(\cdot)$ *occurring in* 3.4(γ) *be bounded on some neighborhood* V_0 *of* 0 *in* Λ.

Then the dilation $S(\cdot)$ *of* $K(\cdot, 0)$ *is strongly continuous on* Λ, *and the dilation* $S(-)^*S(\cdot)$ *of* $K(\cdot, -)$ *is strongly continuous on* $\Lambda \times \Lambda$.

Proof. We appeal to Cor. 3.6 taking $\Gamma = \Lambda$ and therefore $\oplus$ identical to $+$. Since $X(\cdot)$ has a propagator, the conditions in 3.4 prevail; moreover by 3.4(b), $\beta(t) = \sqrt{\gamma(t)}$. Thus (i)–(iv) entail all the premisses of Cor. 3.6. Hence by Cor. 3.6, $S(\cdot)$ is strongly continuous on Λ to $\mathrm{CL}(\mathscr{S}_X, \mathscr{S}_X)$. It follows easily that $S(-)^*$ and $S(-)^*S(\cdot)$ are strongly continuous on Λ and on $\Lambda \times \Lambda$. □

Thm. 5.1 establishes the validity of the Procedure 1.4. For once we have accomplished the hard step (i), i.e. associated with the given function $R(\cdot)$ on Λ to $\mathrm{CL}(W, W)$ satisfying $R(0) = I_W$ the appropriate PD kernel $K(\cdot\cdot)$ on $\Lambda \times \Lambda$ such that $R(\cdot) = K(\cdot, 0)$, Thm. 5.1 assures us that the propagator or controller $S(\cdot)$ of the variety with covariance $K(\cdot\cdot)$ is a dilation of $R(\cdot)$. Cor. 5.2 tells us when for a topological semi-group Λ, this dilation will be strongly continuous.

We shall now deduce improved explicit versions of several dilations theorems from the standpoint of Procedure 1.4 by appeal to 5.1 and 5.2. We begin with the following simplification of Nagy's Principal Theorem in which his premise (c), [26, p. 21], is mitigated; cf. remarks following (4.5):

5.3 THEOREM (simplified Nagy thm.). *Let* (i) Λ *be an involutory* (*additive but not necessary Abelian*) *semi-group with neutral element* 0; (ii) W *be a Hilbert space over* $\mathbb{F}$; (iii) $R(\cdot)$ *be a* PD *function on* Λ *to* $\mathrm{CL}(W, W)$ (cf. 4.1(b)) *such that* $R(0) = I_W$, *and* *having the* (*mild*) *translational property*: $\exists$ *a function* $\gamma(\cdot)$ *on* Λ *to* $\mathbb{R}_{0+} \ni$

$$\forall \lambda, t \in \Lambda, \qquad 0 \leq R(\lambda^* + t^* + t + \lambda) \leq \gamma(t) \cdot R(\lambda^* + \lambda).$$

Then

(a) $\exists$ *a Hilbert space* $\mathfrak{H}_0$ *over* $\mathbb{F}$, $\exists$ *a linear isometry* J *on* W *to* $\mathfrak{H}_0$ *and* $\exists$ *a semi-group* $S(\cdot)$ *on* Λ *to* $\mathrm{CL}(\mathfrak{H}_0, \mathfrak{H}_0) \ni$

$$\forall \lambda \in \Lambda, \qquad R(\lambda) = J^* S(\lambda) J,$$

i.e. $R(\cdot)$ *has the dilation* $S(\cdot)$ *in* $\mathfrak{H}_0$. *Here* $\mathfrak{H}_0 = \mathscr{S}_X$, $J = X(0)$ *and* $S(\cdot)$ *is the propagator of* $X(\cdot)$, *where* $X(\cdot)$ *is a variety on* Λ *having the covariance function* $R(\cdot)$ (cf. 2.10, 4.12);

(b) $\forall \lambda \in \Lambda$, $S(\lambda^*) = S(\lambda)^*$ *and* $|S(\lambda)| \leqq \sqrt{\gamma(\lambda)}$.

Proof. By (iii) and Thm. 4.13, a variety $X(\cdot)$ on Λ to $\mathrm{CL}(W, \mathfrak{H})$ possessing the covariance function $R(\cdot)$ ($\mathfrak{H}$ being as in 2.10) has a propagator $S(\cdot)$ satisfying the conditions in (b). By Thm. 5.1, this $S(\cdot)$ is the dilation of $K(\cdot, 0)$, where $K(\lambda, \lambda') \underset{d}{=} R(\lambda'^* + \lambda)$, i.e. a dilation of $R(\cdot)$ in $\mathscr{S}_X$. Thus (a) holds with $\mathfrak{H}_0 = \mathscr{S}_X \subseteq \mathfrak{H}$ and $J = X(0)$, and so does (b). □

In the same way, starting from Thms. 4.14, 4.15 (instead of 4.13) we arrive at the following dilation theorem for Banach algebra. This is essentially a restatement of Stinespring's [32, Thm. 1] for C^*-algebras:

5.4 THEOREM. *Let* (i) $\mathbb{A}$ *be a Banach algebra over* $\mathbb{F}$ *with unit* $\mathbf{1} \ni |\mathbf{1}| = 1$ *and an isometric involution* $*$, (ii) $R(\cdot)$ *be a linear operator on* (*the vector space*) $\mathbb{A}$ *to* $\mathrm{CL}(W, W)$ *where is a Hilbert space over* $\mathbb{F}$, (iii) $R(\cdot)$ *be a* PD *function on* (*the* *s.g.) $\mathbb{A}$ *to* $\mathrm{CL}(W, W)$ *and* $R(\mathbf{1}) = I_W$. *Then*

(a) $\exists$ *a Hilbert space* $\mathfrak{H}_0$, $\exists$ *a* **-representation* $S(\cdot)$ *of* $\mathbb{A}$ *in* $\mathrm{CL}(\mathfrak{H}_0, \mathfrak{H}_0)$, *and* $\exists$ *a linear isometry* J *on* W *to* $\mathfrak{H}_0$ *such that*

$$\forall a \in \mathbb{A}, \qquad R(a) = J^* S(a) J.$$

This $S(\cdot)$ *is a dilation of* $R(\cdot)$ *in* $\mathfrak{H}_0$.

(b) *In particular,* (a) *holds when* $\mathbb{A}$ *is an Abelian* C^**-algebra with unit* $\mathbf{1}$ *and* $R(\cdot)$ *is any positive linear operator on* $\mathbb{A}$ *to* $\mathrm{CL}(W, W)$.

Proof. (a) By (i)–(iii) and Thm. 4.15, we have

$$(1) \qquad \forall a \in \mathbb{A}, \qquad R(a) = J^* S(a) J$$

where $S(\cdot)$ is a *-representation of $\mathbb{A}$ in $\mathrm{CL}(\mathfrak{H}_0, \mathfrak{H}_0)$ and $J \in \mathrm{CL}(W, \mathfrak{H}_0)$. By 4.15(b), $J = X(\mathbf{1})$, where $X(\cdot)$ is a variety with covariance function $R(\cdot)$. But now

$$X(\mathbf{1})^* X(\mathbf{1}) = R(\mathbf{1}) = I_W,$$

i.e. J is an isometry. Hence by (1), $S(\cdot)$ is a dilation of $R(\cdot)$.

Part (b) follows from (a) in view of Stinespring's [1, Thm. 4]; cf. 4.16, 4.17. □

An important application of Thm. 5.4 is to the Abelian C^*-algebra $\mathbb{A} = C(X; \mathbb{C})$, of complex-valued continuous functions on a compact subset X of a Hausdorff space (sup norm). It transpires that for certain (unclosed) subalgebras $\mathbb{D}$ of $\mathbb{A}$, linear operators on $\mathbb{D}$ to $\mathrm{Cl}(W, W)$ of a certain type extend to positive linear operators $R(\cdot)$ on $\mathbb{A}$. The application of 5.4(b) to this extension $R(\cdot)$ on $\mathbb{A}$ often yields a valuable dilation theorem for the original operator on $\mathbb{D}$. This observation and the revelation of its scope are due to Arveson [5, 0.1, 1.22 et seq.], who are able to relate the Silov boundary of X relative to $\mathbb{D}$ to his concept of the "support" of the resulting dilation $S(\cdot)$. We shall only mention here the pioneering theorem of Lebow [16, p. 84], which emerges as an easy corollary of Arveson's theorem, but which seems to have been seminal to this research.

5.5 THEOREM (Lebow). *Let* (i) $T \in \mathrm{CL}(W, W)$, *where* W *is a Hilbert space over* $\mathbb{C}$, (ii) *the* (*necessarily compact*) *spectral set* $X \subset \mathbb{C}$ *of* T *be such that the algebra* $\mathbb{D}$ *of rational* *functions on $\mathbb{C}$ with poles in $\mathbb{C}\backslash X$ is a Dirichlet algebra. Then*

(a) *$\exists$ a Hilbert space $\mathfrak{H}$, $\exists$ an isometry J on W to $\mathfrak{H}$, and $\exists$ a normal operator $\tilde{T} \in \mathrm{CL}(\mathfrak{H}, \mathfrak{H})$ such that*

$$\forall f \in \mathbb{D}, \qquad f(\tilde{T}) = J^* f(T) J.$$

(b) *$\sigma(\tilde{T}) \subseteq \partial X$, where $\sigma(\tilde{T}) \underset{d}{=}$ the spectrum of $\tilde{T}$.*

Here the fact of X being a spectral set of T and the maximum modulus principle yield

$$\forall f \in \mathbb{D}, \qquad |f(T)| \leqq \max_{z \in X} |f(z)| = \max_{z \in \partial X} |f(z)|.$$

This together with the fact that $\mathbb{D}$ is a Dirichlet algebra enables us to obtain by extension a positive linear operator $R(\cdot)$ on $\mathbb{A} \underset{d}{=} C(X, \mathbb{C})$ to $\mathrm{CL}(W, W)$ such that

$$(1) \qquad \forall f \in \mathbb{D}, \qquad f(T) = R(f).$$

The application of 5.4(b) to $R(\cdot)$ now yields

$$(2) \qquad \forall f \in \mathbb{A}, \qquad R(f) = J^* S(f) J.$$

Since $\mathbb{A}$ is Abelian, each f in $\mathbb{A}$ is "normal" and so therefore is $S(f)$; cf. 4.11(a). The desired equality follows from (1) and (2) on letting $\tilde{T} \underset{d}{=} S(\mathbf{1})$ and remembering that the propagator $S(\cdot)$ is a multiplicative s.g. The relation between $\sigma(\tilde{T})$ and ∂X follows from Arveson's relation between the "support" of $S(\cdot)$ and the Silov boundary of X relative to $\mathbb{D}$.

It follows from a classical theorem of J. L. Walsh that the $\mathbb{D}$ in 5.5 is a Dirichlet algebra when $\mathbb{C}\backslash X$ is connected; cf. [16, p. 66]. Lebow's Thm. in this special case embraces several classical dilation theorems; cf. [16, p. 86].

From our standpoint the pioneering work of Stinespring, Lebow and Arveson as well as some of the earlier work of Halmos and Nagy falls under step (i) of the Procedure 1.4: it is a quest for significant PD kernels. Their researches tell us that kernels which are derivable from positive or completely positive operators on a C^*-algebra, or ones whose extensions are so derivable, are PD. This knowledge obviates the need for ab initio demonstration of positive-definiteness in individual cases. We should point out, however, that not all situations are reasily amenable to the Lebow–Arveson algebraic approach. A recalcitrant instance is the strongly continuous contractive semi-group $(C_t\colon t \in \mathbb{R}_{0+})$. But as Nagy showed, cf. [26, p. 32] and [27, p. 30], it generates a stationary PD kernel, and Procedure 1.4 works.

We turn next to the dilation of a W-to-W nonnegative hermitian operator-valued measure, where W is a Hilbert space (Naimark's Thm.). Step (i) of Procedure 1.4, to determine the PD kernel corresponding to such a measure, is accomplished in the following lemma, valid for a W-to-W^* operator-valued measure of this type, W being any Banach space.

5.6 LEMMA. *Let* (i) *$\mathcal{P}$ be a pre-ring over a space Ω,* (ii) *W be a Banach space over $\mathbb{F}$,* (iii) *$M(\cdot)$ be a W-to-W^* nonnegative hermitian operator-valued, finitely additive, measure on $\mathcal{P}$,* (iv) *$\forall A, B \in \mathcal{P}$, $K(A, B) \underset{d}{=} M(A \cap B)$. Then $K(\cdot\,\cdot)$ is a PD kernel on $\mathcal{P} \times \mathcal{P}$ to $\mathrm{CL}(W, W^*)$.*

Proof. Let $r \in \mathbb{N}_+$, $C_1, \cdots, C_r \in \mathrm{CL}(W, W)$ and $A_1, \cdots, A_r \in \mathcal{P}$. We must first show that

$$(\mathrm{I}) \qquad \sum_{i=1}^{r} \sum_{j=1}^{r} C_j^* K(A_i, A_j) C_i \geq 0.$$

Now to the sets $A_1, \cdots, A_r$ in $\mathcal{P}$ correspond disjoint sets $B_1, \ldots, B_q$ in $\mathcal{P}$, where $q \in \mathbb{N}_+$, such that we have the partitioning

$$(1) \qquad A_i = \bigcup_{n \in N_i} B_n, \qquad \{1, \cdots, q\} = \bigcup_{i=1}^{r} N_i.$$

It follows easily that

$$A_i \cap A_j = \bigcup_{n \in N_i \cap N_j} B_n, \quad \text{with } B_n \text{ disjoint},$$

and therefore from the finite-additivity of $M(\cdot)$,

$$K(A_i, A_j) = \sum_{n \in N_i \cap N_j} M(B_n) = \sum_{n=1}^{q} M(B_n)\chi_{N_i \cap N_j}(n),$$

where χ_S stands for the indicator of set S. Thus

$$\begin{aligned} \text{LHS(I)} &= \sum_{i=1} \sum_{j=1} C_j^* \left\{ \sum_{n=1}^{q} M(B_n)\chi_{N_i}(n)\chi_{N_j}(n) \right\} C_i \\ &= \sum_{n=1}^{q} \left\{ \sum_{j=1}^{r} C_j^* \chi_{N_j}(n) \right\} \cdot M(B_n) \cdot \left\{ \sum_{i=1}^{r} C_1 \chi_{N_i}(n) \right\} \\ &= \sum_{n=1}^{q} C(n)^* M(B_n) C(n) \geq 0, \quad \text{by } 2.4, \end{aligned}$$

since $C(n) \underset{d}{=} \sum_{i=1}^{r} C_i \chi_{N_i}(n) \in \text{CL}(W, W)$.

This establishes (I), i.e. the first condition in Def. 2.5. The second is easily checked, since $K(A, B) = K(B, A)$ is hermitian. □

Now let $\mathcal{P}$, W, $M(\cdot)$ and $K(\cdot\cdot)$ be as in Lemma 5.6 and let $\mathfrak{H}$ be as in the Kernel Thm. 2.10, and $T(\cdot)$ on $\mathcal{P}$ to $\text{CL}(W, \mathfrak{H})$ be the variety with covariance kernel $K(\cdot\cdot)$, so that

$$(5.7) \qquad \forall A, B \in \mathcal{P}, \qquad T(B)^* T(A) = M(A \cap B).$$

We have studied this set function $T(\cdot)$ extensively in [21] in the case where W is a Hilbert space and the measure $M(\cdot)$ is strongly countably additive (s.c.a.) on $\mathcal{P}$. We call $T(\cdot)$ a *W-to-$\mathfrak{H}$ countably additive quasi-isometric (c.a.q.i.) measure on $\mathcal{P}$, with control measure* $M(\cdot)$ because (5.7) entails that

$$\forall A \in \mathcal{P}, \qquad \text{cls.}[T(A)\{\sqrt{M(A)}\}^{\sim 1}] \text{ is a partial isometry},$$

the superscript ~ 1 indicating the generalized inverse, and also entails that for an s.c.a. measure $M(\cdot)$ on $\mathcal{P}$, $T(\cdot)$ is s.c.a. on $\mathcal{P}$. Thus for a Hilbert space W and an s.c.a. measure $M(\cdot)$, *the variety having the* $K(\cdot\cdot)$ *of* 5.6 *as covariance kernel is a W-to-$\mathfrak{H}$ c.a.q.i. measure on $\mathcal{P}$ with control measure* $M(\cdot)$. This completes step (ii) of Procedure 1.4, for in the Naimark dilation theorem $M(\cdot)$ is given to be s.c.a. on a σ-algebra $\mathcal{B}$ over Ω and $M(\Omega) = I_W$.

The situation just described comes within the scope of Thm. 5.3, since even a *pre-algebra over Ω is an Abelian semi-group under $\cap$ with neutral element Ω, and therefore a* *s.g. *with* $A^* = A$, and $M(\cdot)$ has the translational property 5.3(iii) as is easy to check. By Thm. 5.3 the variety $T(\cdot)$ on $\mathcal{B}$ has a propagator. But more insight into its nature is provided by the theory of c.a.q.i. measures. To see this, let $\mathcal{D}$ be just a δ-ring over Ω, and the W-to-W nonnegative hermitian measure $M(\cdot)$ be s.c.a. on $\mathcal{D}$, W now being a Hilbert space. Let also

$$(5.8) \quad \begin{cases} \mathcal{D}^{\text{loc}} \underset{d}{=} \{B : B \subseteq \Omega \text{ and } \forall \Delta \in \mathcal{D}, B \cap \Delta \in \mathcal{D}\}, \\ \forall B \in \mathcal{D}^{\text{loc}}, \qquad \mathcal{M}_T(B) \underset{d}{=} \mathfrak{S}\{T(\Delta)(W) : \Delta \in \mathcal{D} \cap 2^B\} \subseteq \mathfrak{H}, \\ \forall B \in \mathcal{D}^{\text{loc}}, \qquad Q_T(B) \underset{d}{=} \text{the orthogonal projection on } \mathcal{M}_T(\Omega) \text{ onto } \mathcal{M}_T(B). \end{cases}$$

We then have the following theorem:

5.9 THEOREM. (a) *$\mathcal{M}_T(\cdot)$ is a countably-additive, orthogonally scattered (c.a.o.s.) subspace-valued measure on the σ-algebra $\mathcal{D}^{\text{loc}}$.* (b) *$Q_T(\cdot)$ is a s.c.a. projection-valued measure on $\mathcal{D}^{\text{loc}}$ for the Hilbert subspace $\mathcal{M}_T(\Omega)$ of $\mathfrak{H}$.*

Proof. The proof given in [21, 10.25] is correct, but needlessly involves integration. A clearer, integration-free proof (which should have appeared in [21, § 8]) is given in Appendix H. □

5.10 DEFINITION. Let $T(\cdot)$ be a W-to-$\mathfrak{H}$ c.a.q.i. measure on a δ-ring $\mathcal{D}$, and $\mathcal{M}_T(\cdot)$, $Q_T(\cdot)$ be as in (5.8). Then

(a) $\mathcal{M}_T(\cdot)$ is called the *spatial measure* of $T(\cdot)$,

(b) $Q_T(\cdot)$ is called the *spatial spectral measure* of $T(\cdot)$.

For dilation theory we need besides Thm. 5.9(b) the relation:

$$(5.11) \qquad \forall\Delta \in \mathcal{D} \text{ and } \forall B \in \mathcal{D}^{\text{loc}}, \qquad Q_T(B)\cdot T(\Delta) = T(\Delta \cap B).$$

This is immediate from the relations

$$T(\Delta)(W) = T(\Delta \cap B)(W) + T(\Delta\backslash B)(W)$$

$$T(\Delta \cap B)(W) \perp T(\Delta\backslash B)(W).$$

Since $\cap$ is the "+ operation" of the semi-group $\mathcal{D}$, (5.11) shows that *the spatial spectral measure $Q_T(\cdot)$ is the controller of $T(\cdot)$* in the sense of Def. 3.2(b).[17]

Now let $\mathcal{D} = \mathcal{B}$ be a σ-algebra and $M(\Omega) = I_W$. Then $\mathcal{D}^{\text{loc}} = \mathcal{B}$, and all the premises of Thm. 5.1 are satisfied with $\Lambda = \mathcal{B}$. We thus immediately get the following explicit version of the Naimark theorem [26, p. 6]:

5.12 NAIMARK'S THEOREM (explicit form). *Let*

(i) *W be a Hilbert space over $\mathbb{F}$,*

(ii) *$M(\cdot)$ be a W-to-W nonnegative hermitian operator-valued s.c.a. measure on a σ-algebra $\mathcal{B}$ over Ω such that $M(\Omega) = I_W$,*

(iii) *the* (PD) *kernel $K(\cdot\,\cdot)$ on $\mathcal{B} \times \mathcal{B}$ be defined by*

$$\forall A, B \in \mathcal{B}, \qquad K(A, B) \underset{d}{=} M(A \cap B),$$

and $\mathfrak{H}$ be as in the Kernel Thm. 2.10 *for this $K(\cdot\,\cdot)$,*

(iv) *$T(\cdot)$ be the W-to-$\mathfrak{H}$ c.a.q.i. measure on $\mathcal{B}$ with control measure $M(\cdot)$, i.e. the variety with covariance kernel $K(\cdot\,\cdot)$,*

(v) *$Q_T(\cdot)$ be the spatial spectral measure of $T(\cdot)$ on $\mathcal{B}$* (cf. Def. 5.10).

Then

$$\forall B \in \mathcal{B}, \qquad M(B) = T(\Omega)^* Q_T(B) T(\Omega).$$

Hence $Q_T(\cdot)$ is a dilation of $M(\cdot)$.

5.13 *Remarks.* Thm. 5.12 shows that the dilation space for the operator-valued measure $M(\cdot)$ is the subspace $\mathcal{M}_T(\Omega)$, i.e. $\mathcal{S}_T$, of the Hilbert space $\mathfrak{H}$ given by Thm. 2.10 for the kernel $K(\cdot\,\cdot)$ arising from $M(\cdot)$. As is clear from Thm. 2.10 the corresponding subspace of any other Hilbert space $\mathfrak{K}$ of equal or greater dimension than $\mathfrak{H}$ can also be taken as the dilation space of $M(\cdot)$. One such space $\mathfrak{K}$ is

$$\bar{\mathcal{L}}_{2,W} = \text{the completion of } \mathcal{L}_{2,W}, \qquad \mathcal{L}_{2,W} \underset{d}{=} L_2(\Omega, \mathcal{B}, M;\, W).$$

[17] Here the term *controller* is more appropriate than *propagator*, since the semi-group $\mathcal{P}$ does not represent time or space-time or phase space.

For the definition of $\mathscr{L}_{2,W}$ see [21: § 9].[18] For the Hilbert space $\mathfrak{K} = \bar{\mathscr{L}}_{2,W}$, the Hilbertian variety we get is the *indicator c.a.q.i. measure* $M_\chi(\cdot)$, defined in [21, 10.13]. The dilation of $M(\cdot)$, i.e. the spatial spectral measure of $M_\chi(\cdot)$, now turns out to be the most familiar spectral measure over L_2 spaces, to wit multiplication by the indicator functions of set in $\mathscr{B}$.

6. Image-extensions. In this section we shall deal with image-extensions in the sense of Def. 1.5. In our scheme in which projections give way to isometries, this concept corresponds to the classical one of extension to a larger Hilbert space (as opposed to dilation). Our purpose is to show that our simplified version 5.3 of the Nagy Principal Theorem, with only the mild translational requirement, is potent enough to yield the improved (translation-requirement-free) version of the Halmos theorem on normal extensions due to Bram [6, Thm. 1]. Apart from this one aspect, our proof follows the one due to Nagy [26, § 10]. Like Nagy we need the following lemma, the proof of which we leave to the reader; cf. [26, p. 20].

6.1 LEMMA. *Let* (i) *W and $\mathfrak{H}$ be Hilbert spaces over $\mathbb{F}$,* (ii) *J be a linear isometry on W to $\mathfrak{H}$,* (iii) *$T \in \mathrm{CL}(W, W)$ and $\tilde{T} \in \mathrm{CL}(\mathfrak{H}, \mathfrak{H})$. Then the following conditions are equivalent*:

(α) $JTJ^{-1} \subseteq \tilde{T}$,

(β) $T = J^{-1}\tilde{T}J$ *and* $\tilde{T}(\mathscr{R}_J) \subseteq \mathscr{R}_J$, $\mathscr{R}_J \underset{d}{=}$ *Range of J,*

(γ) $T = J^*\tilde{T}J$ *and* $T^*T = J^*\tilde{T}^*\tilde{T}J$.

6.2 THEOREM *Let* (i) *Λ be an (additive) Abelian semi-group with neutral element* 0; (ii) *$\forall \lambda \in \Lambda$, $T_\lambda \in \mathrm{CL}(W, W)$ where W is a Hilbert space over $\mathbb{F}$. Then the following conditions are equivalent*:

(α) *$(T_\lambda : \lambda \in \Lambda)$ is a semi-group in $\mathrm{CL}(W, W)$, and $\forall$ functions $w(\cdot)$ on Λ to W, $\forall r \in \mathbb{N}$ and $\forall \lambda_1, \cdots, \lambda_r \in \Lambda$,*

$$\sum_{i=1}^{r} \sum_{j=1}^{r} (T_{\lambda_i}\{w(\lambda_j)\}, T_{\lambda_j}\{w(\lambda_i)\}) \geqq 0,$$

(β) *$\exists$ a Hilbert space $\mathfrak{H}_0$ over $\mathbb{F}$, $\exists$ a linear isometry J on W into $\mathfrak{H}_0$ and $\exists$ a semi-group $(\tilde{T}_\lambda : \lambda \in \Lambda)$ of continuous normal operators on $\mathfrak{H}_0$ to $\mathfrak{H}_0$ such that*

$$\forall \lambda \in \Lambda, \quad JT_\lambda J^{-1} \subseteq \tilde{T}_\lambda,$$

i.e. the normal operator $\tilde{T}_\lambda$ is an extension of the image $JT_\lambda J^{-1}$ in $\mathrm{CL}(\mathscr{R}_J, \mathscr{R}_J)$ of T_λ in $\mathrm{CL}(W, W)$, $\mathscr{R}_J \underset{d}{=}$ range of J.

Proof. Let (α) hold. Then, following Nagy, define + and * in $\Lambda^2 \underset{d}{=} \Lambda \times \Lambda$ and $R(\cdot)$ on Λ^2 by

$$(1) \qquad (\lambda_1; \lambda_2) + (\lambda_1'; \lambda_2') \underset{d}{=} (\lambda_1 + \lambda_1'; \lambda_2 + \lambda_2'), \qquad (\lambda_1, \lambda_2)^* \underset{d}{=} (\lambda_2, \lambda_1),$$

$$(2) \qquad R(\lambda_1; \lambda_2) \underset{d}{=} T_{\lambda_2}^* T_{\lambda_1}.$$

Obviously,

(3) Λ^2 is a *s.g. with neutral element (0, 0),

and exactly as in [26, p. 36] we deduce from the inequality in (α) that

(4) $R(\cdot)$ is a PD function on Λ^2 to $\mathrm{CL}(W, W)$.

[18] $\mathscr{L}_{2,W}$ is always an inner product space. Only for so-called "adequate" measures $M(\cdot)$, is $\mathscr{L}_{2,W}$ complete [21, 9.9]. For important further work on this question see Mandrekar and Salehi [24].

We now claim that $R(\cdot)$ fulfills the mild translational requirement of Thm. 5.3, viz.

$$(5)\quad \begin{cases} \exists \text{ a function } \gamma(\cdot) \text{ on } \Lambda^2 \text{ to } \mathbb{R}_{0+} \ni \forall(\lambda_1;\lambda_2) \text{ and } (t_1;t_2)\in\Lambda^2, \\ R\{(\lambda_1;\lambda_2)^*+(t_1;t_2)^*+(t_1;t_2)+(\lambda_1;\lambda_2)\} \\ \qquad \le \gamma(t_1;t_2)\cdot R\{(\lambda_1;\lambda_2)^*+(\lambda_1;\lambda_2)\}. \end{cases}$$

Proof of (5). By (1) and the commutativity of + in Λ^2, the argument of $R(\cdot)$ on the LHS of (5) is

$$(t_1+t_2+\lambda_1+\lambda_2;\, t_1+t_2+\lambda_1+\lambda_2).$$

Hence by (2) and the semi-group properties of $(T_\lambda:\lambda\in\Lambda)$,

$$\begin{aligned}\text{LHS}(5) &= T^*_{t_1+t_2+\lambda_1+\lambda_2}\cdot T_{t_1+t_2+\lambda_1+\lambda_2} = T^*_{\lambda_1+\lambda_2}T^*_{t_1+t_2}T_{t_1+t_2}T_{\lambda_1+\lambda_2} \\ &\le |T^*_{t_1+t_2}T_{t_1+t_2}|\cdot T^*_{\lambda_1+\lambda_2}T_{\lambda_1+\lambda_2},\end{aligned}$$

the last step stemming from the triviality that for $A, H\in \mathrm{CL}(W, W)$, $0\le H \Rightarrow 0\le A^*HA \le |H|A^*A$. By letting $\gamma(t_1;t_2) \underset{d}{=} |T_{t_1+t_2}|^2$, it thus follows that

$$\text{LHS}(5) \le \gamma(t_1;t_2)\cdot R(\lambda_1+\lambda_2;\lambda_1+\lambda_2).$$

This proves (5).

By (3)–(5) and Thm. 5.3(a), $\exists$ a Hilbert space $\mathfrak{H}_0$, $\exists$ an isometry J on W to $\mathfrak{H}_0$ and $\exists$ a s.g. $S(\cdot)$ on Λ^2 to $\mathrm{CL}(\mathfrak{H}_0,\mathfrak{H}_0)$ such that

$$(6)\qquad \forall(\lambda_1;\lambda_2)\in\Lambda^2,\qquad T^*_{\lambda_2}T_{\lambda_1} \underset{d}{=} R(\lambda_1;\lambda_2) = J^*S(\lambda_1;\lambda_2)J.$$

Since Λ^2 is Abelian, each $(\lambda_1;\lambda_2)$ is "normal" and so therefore is each $S(\lambda_1;\lambda_2)$; cf. 4.11. By letting

$$(7)\qquad \forall\lambda\in\Lambda,\qquad \tilde T_\lambda \underset{d}{=} S(\lambda;0),$$

it follows at once that

$$(8)\qquad (\tilde T_\lambda:\lambda\in\Lambda) \text{ is a s.g. of continuous normal operators on } \mathfrak{H}_0 \text{ to } \mathfrak{H}_0.$$

Also from 5.3(b) and (7),

$$S(\lambda_1;\lambda_2) = S\{(\lambda_2;0)^*+(\lambda_1;0)\} = \{S(\lambda_2;0)\}^*\cdot S(\lambda_1;0) = \tilde T^*_{\lambda_2}\tilde T_{\lambda_1}.$$

Thus (6) can be restated:

$$(9)\qquad \forall\lambda,\lambda'\in\Lambda,\qquad T^*_{\lambda'}T_\lambda = J^*\tilde T^*_{\lambda'}\tilde T_\lambda J.$$

Taking $\lambda'=0$, we get

$$(10)\qquad \forall\lambda\in\Lambda,\qquad T_\lambda = J^*\tilde T_\lambda J.$$

By (9), (10) and the last lemma,

$$(11)\qquad JT_\lambda J^{-1}\subseteq \tilde T_\lambda.$$

By (8) and (11) we have (β).

Next let (β) hold. Then $\forall\lambda\in\Lambda$, $JT_\lambda J^{-1}\subseteq\tilde T_\lambda$. Hence by the last lemma,

$$(12)\qquad \forall\lambda\in\Lambda,\quad T_\lambda = J^{-1}\tilde T_\lambda J = J^*\tilde T_\lambda J,\quad\text{and}\quad \tilde T_\lambda(\mathscr{R}_J)\subseteq\mathscr{R}_J.$$

From (12) and the fact that $(\tilde T_\lambda:\lambda\in\Lambda)$ is a s.g., it is easy to deduce that

$$(13)\qquad (T_\lambda:\lambda\in\Lambda) \text{ is a s.g. in } \mathrm{CL}(W, W).$$

Also, $\forall w, w' \in W$,

$$(T_\lambda w', T_{\lambda'} w)_W = (J^{-1}\tilde{T}_\lambda Jw', J^{-1}\tilde{T}_{\lambda'} Jw)_W \quad \text{by (12)}$$
$$= (\tilde{T}_\lambda Jw', \tilde{T}_{\lambda'} Jw)_{\mathfrak{H}_0}, \quad \text{as } J^{-1} = \text{isometry}$$
$$= (\tilde{T}^*_{\lambda'} Jw', \tilde{T}^*_\lambda Jw)_{\mathfrak{H}_0}, \tag{14}$$

the last step being a consequence of Fuglede's thm., which guarantees that $\tilde{T}_\lambda$ and $\tilde{T}^*_{\lambda'}$ commute, since, by (β), $\tilde{T}_\lambda$ and $\tilde{T}_{\lambda'}$ do. It follows from (14) that

$$\sum_{i=1}^{r}\sum_{j=1}^{r} (T_{\lambda_i}\{w(\lambda_j)\}, T_{\lambda_j}\{w(\lambda_i)\}) = \sum_{i=1}^{r}\sum_{j=1}^{r} (\tilde{T}^*_{\lambda_j} J\{w(\lambda_j)\}, \tilde{T}^*_{\lambda_i} J\{w(\lambda_i)\})$$
$$= \left|\sum_{i=1}^{r} \tilde{T}^*_{\lambda_i} J\{w(\lambda_i)\}\right|^2 \geqq 0. \tag{15}$$

By (13) and (15) we have (α). □

It is worth restating Thm. 6.2 in terms of subnormality. Let $(T_\lambda : \lambda \in \Lambda)$ be a semi-group of subnormal operators on a Hilbert space W; cf. [6, Def. p. 75]. Then for each λ in Λ, there exists a Hilbert space $\mathfrak{H}_\lambda$ and a continuous normal operator $\tilde{T}_\lambda$ thereon such that $W \subseteq \mathfrak{H}_\lambda$ and $T_\lambda \subseteq \tilde{T}_\lambda$. But the spaces $\mathfrak{H}_\lambda$ need not be equal, nor when they are, need the $\tilde{T}_\lambda$ form a semi-group. This suggests a definition to cover the exceptional case:

6.3 Definition. Let Λ, W, T_λ be as in 6.2(i)(ii). We say that the family $(T_\lambda : \lambda \in \Lambda)$ is *semi-group-subnormal* iff ∃ a Hilbert space $\mathfrak{H}_0$ and ∃ a semi-group $(\tilde{T}_\lambda : \lambda \in \Lambda)$ of normal operators in $\mathrm{CL}(\mathfrak{H}_0, \mathfrak{H}_0)$ such that $W \subseteq \mathfrak{H}_0$ and $\forall \lambda \in \Lambda$, $T_\lambda \subseteq \tilde{T}_\lambda$.

It is clear that if in 6.2(β) we "identify" W and $\mathscr{R}_J$ (thereby obtaining $W \subseteq \mathfrak{H}_0$ and $T_\lambda \subseteq \tilde{T}_\lambda$), then we can restate Thm. 6.2 more briefly as follows

6.4 Theorem. *Let* (i), (ii) *be as in* 6.2. *Then* $(T_\lambda : \lambda \in \Lambda)$ *is semi-group-subnormal iff.* 6.2(α) *holds.*

Now for single operators T, $\tilde{T}$, the statements "$T \in \mathrm{CL}(W, W)$" and "$(T^n : n \in \mathbb{N}_{0+})$ is a s.g. in $\mathrm{CL}(W, W)$" are equivalent, as are the statements "$\tilde{T} \in \mathrm{CL}(\mathfrak{H}_0, \mathfrak{H}_0)$ is normal" and "$(\tilde{T}^n : n \in \mathbb{N}_{0+})$ is a s.g. of normal operators in $\mathrm{CL}(\mathfrak{H}_0, \mathfrak{H}_0)$". Also, of course, $T \subseteq \tilde{T} \Rightarrow \forall n \in \mathbb{N}_{0+}$, $T^n \subseteq \tilde{T}^n$. Hence for $\Lambda = \mathbb{N}_{0+}$, Thm. 6.4 reduces to the assertion that $T \in \mathrm{CL}(W, W)$ *is subnormal iff* $\forall r \in \mathbb{N}_+$, $\forall k_1, \cdots, k_r \in \mathbb{N}_{0+}$ and $\forall w_1, \cdots, w_r \in W$,

$$\sum_{i=1}^{r}\sum_{j=1}^{r} (T^{k_i}(w_j), T^{k_j}(w_i)) \geqq 0.$$

This is a restatement of Bram's Thm. 1; cf. [6, p. 77]. Its derivation shows that *Thm.* 6.2 *constitutes in our new setting an extension to arbitrary Abelian semi-groups of the Bram theorem.*

Appendix A. Proof of Main Lemma. 2.6. We shall show that $(\alpha) \Leftrightarrow (\beta)$ and $(\beta) \Leftrightarrow (\gamma)$. The abbreviations $w_i = w(\lambda_i)$ and $C_i = C(\lambda_i)$ will be tacitly understood in what follows.

Let (α) hold. Only the first conditon in (β) has to be proved. Let $w(\cdot)$ be on Λ to W, $r \in \mathbb{N}_+$ and $\lambda_1, \cdots, \lambda_r \in \Lambda$. Take any $f_1, \cdots, f_r \in W' \backslash \{0\}$, W' being the dual of W. Then[19]

$$\exists w_0 \in W \ni \forall j = 1, \cdots, r, \qquad f_j(w_0) \neq 0.$$

[19] The quickest way to see this is to observe that the r hyperplanes $N_j \underset{d}{=}$ null space f_j have void interiors, and hence by the Baire theorem their union cannot be equal to W.

Since $f_j \in \mathrm{CL}(W, \mathbb{F})$, therfore obviously

$$C_j(\cdot) \underset{d}{=} \frac{f_j(\cdot)}{f_j(w_0)} w_j \in \mathrm{CL}(W, W) \quad \text{and} \quad C_j(w_0) = w_j.$$

It follows that

$$\begin{aligned} LHS(\beta) &= \sum_{i=1}^{r} \sum_{j=1}^{r} [K(\lambda_i, \lambda_j) C_i(w_0)](C_j(w_0)) \\ &= \left[\left\{\sum_{i=1}^{r} \sum_{j=1}^{r} C_j^* K(\lambda_i, \lambda_j) C_i\right\}(w_0)\right](w_0), \quad \text{cf. 2.2(c)} \\ &\geqq 0 \quad \text{by } (\alpha) \text{ and Def. 2.3(b).} \end{aligned}$$

Thus (β) holds.

Next let (β) hold. Only the first condition in Def. 2.5 has to be proved. Let $C(\cdot)$ be on Λ to $\mathrm{CL}(W, W)$, $r \in \mathbb{N}_+$ and $\lambda_1, \cdots, \lambda_r \in \Lambda$. Then obviously

$$H \underset{d}{=} \sum_{i=1}^{r} \sum_{j=1}^{r} C_j^* K(\lambda_i, \lambda_j) C_i \in \mathrm{CL}(W, W^*).$$

Now let $w, w' \in W$ and $w_i \underset{d}{=} C_i(w)$, $w_i' \underset{d}{=} C_i(w')$. Then

$$(1) \qquad \begin{aligned} [H(w)](w') &= \sum_{i=1}^{r} \sum_{j=1}^{r} [C_j^* K(\lambda_i, \lambda_j) C_i(w)](w') \\ &= \sum_{i=1}^{r} \sum_{j=1}^{r} [K(\lambda_i, \lambda_j)(w_i)](w_j'). \end{aligned}$$

Similarly

$$(2) \qquad [H(w')](w) = \sum_{i=1}^{r} \sum_{j=1}^{r} [K(\lambda_i, \lambda_j)(w_j')](w_i).$$

By (1), (2) and the second condition in (β), it follows that $[H(w)](w') = \overline{[H(w')](w)}$, i.e. H is hermitian. Also by (1) and the first condition in (β), $[H(w)](w) \geqq 0$, i.e. H is nonnegative. This yields the first conditon in Def. 2.5. Thus (α) is proved.

To turn to (β) and (γ), let $\mathcal{U} = \Lambda \times W$, $\alpha_i = (\lambda_i, w_i) \in \mathcal{U}$, $c_i \in \mathbb{F}$, for $i = 1, \cdots, r$. Then by the definition of $k(\cdot\,\cdot)$,

$$(3) \qquad k(\alpha_i, \alpha_j) = [K(\lambda_i, \lambda_j)(w_i)](w_j).$$

Since $K(\lambda_i, \lambda_j)(w_i)$ is semi-linear on W, we also get

$$(4) \qquad \begin{aligned} \sum_{i=1}^{r} \sum_{j=1}^{r} \bar{c}_j k(\alpha_i, \alpha_j) c_i &= \sum_{i=1}^{r} \sum_{j=1}^{r} \bar{c}_j c_i [K(\lambda_i, \lambda_j)(w_i)](w_j) \\ &= \sum_{i=1}^{r} \sum_{j=1}^{r} [K(\lambda_i, \lambda_j)(c_i w_i)](c_j w_j). \end{aligned}$$

Now let (β) hold. Then the first condition therein entails that RHS(4) $\geqq 0$, and so LHS(4) $\geqq 0$. Also, the second condition in (β) implies via (3) the conjugate symmetry of $k(\cdot\,\cdot)$ on $\mathcal{U} \times \mathcal{U}$. Thus ($\gamma$) holds.

Finally, let (γ) hold. Then LHS(4) $\geqq 0$ and so RHS(4) $\geqq 0$. Setting $c_1 = \cdots = c_r = 1$, we have the first condition in (β). The second condition in (β) follows by (3) from the conjugate symmetry of $k(\cdot\,\cdot)$ on $\mathcal{U} \times \mathcal{U}$. Thus ($\beta$) is proved. □

Appendix B. Proof of Congruence Theorem 2.9. Let $\mathcal{U} \underset{d}{=} \Lambda \times W$, and

$$\text{(1)} \qquad \forall(\lambda, w) \in \mathcal{U}, \quad x(\lambda, w) \underset{d}{=} X(\lambda)w, \quad y(\lambda, w) \underset{d}{=} Y(\lambda)w.$$

Then $x(\cdot)$, $y(\cdot)$ are functions on $\mathcal{U}$ to $\mathfrak{H}$, $\mathfrak{K}$, respectively, and obviously

$$\text{(2)} \qquad \mathcal{S}_x = \mathcal{S}_X \subseteq \mathfrak{H}, \qquad \mathcal{S}_y = \mathcal{S}_Y \subseteq \mathfrak{K}.$$

Also by (1), 2.2(c) and (ii), $\forall \alpha = (\lambda, w) \in \mathcal{U}$ and $\forall \alpha' = (\lambda', w') \in \mathcal{U}$,

$$\begin{aligned}(x(\alpha), x(\alpha'))_{\mathfrak{H}} &= (X(\lambda)w, X(\lambda')w')_{\mathfrak{H}} = [\{X(\lambda')^* X(\lambda)\}(w)](w') \\ &= [\{Y(\lambda')^* Y(\lambda)\}(w)](w') = (Y(\lambda)w, Y(\lambda')w')_{\mathfrak{K}} \\ \text{(3)} \qquad &= (y(\alpha), y(\alpha'))_{\mathfrak{K}}.\end{aligned}$$

By (3) and a known result, cf. Parzen [29, p. 472], $\exists$ a unitary operator V on $\mathcal{S}_x$ onto $\mathcal{S}_y$ such that

$$\text{(4)} \qquad \forall \alpha \in \mathcal{U}, \qquad y(\alpha) = V\{x(\alpha)\}.$$

It follows readily from (1), (2) and (4) that this V meets all our demands. □

Appendix C. Proof of Kernel Theorem 2.10. (a) Recall the basic properties of the reproducing kernel Hilbert space $\mathfrak{K}$ of a PD kernel $k(\cdot\,\cdot)$ on $\mathcal{U} \times \mathcal{U}$ to $\mathbb{F}$, where $\mathcal{U}$ is any parameter space. We know, cf. Aronszajn [3, p. 343–345],

$$\text{(1)} \quad \begin{cases} \mathfrak{K} \subseteq \mathbb{F}^{\mathcal{U}}, \\ \{k(\alpha, \cdot): \alpha \in \mathcal{U}\} \text{ is a fundamental subset of } \mathfrak{K}, \\ \forall f \in \mathfrak{K} \text{ and } \forall \alpha \in \mathcal{U}, \qquad (f, k(\alpha, \cdot))_{\mathfrak{K}} = f(\alpha) \in \mathbb{F}, \\ \forall \alpha_1, \alpha_2 \in \mathcal{U}, \qquad (k(\alpha_1, \cdot), k(\alpha_2, \cdot))_{\mathfrak{K}} = k(\alpha_1, \alpha_2), \\ \forall \alpha \in \mathcal{U}, \qquad |k(\alpha, \cdot)|^2_{\mathfrak{K}} = k(\alpha, \alpha). \end{cases}$$

Now, cf. 2.6(γ), we take for $k(\cdot\,\cdot)$ the $\mathbb{F}$-valued PD kernel on $(\Lambda \times W) \times (\Lambda \times W)$ given by

$$\text{(C.1)} \qquad k\{(\lambda_1, w_1), (\lambda_2, w_2)\} = [K(\lambda_1, \lambda_2)(w_1)](w_2) \in \mathbb{F}.$$

Here $\mathcal{U} = \Lambda \times W$. Next we define $\forall \lambda_1 \in \Lambda$ and $\forall w_1 \in W$,

$$\text{(C.2)} \qquad T(\lambda_1)(w_1) \underset{d}{=} k\{(\lambda_1, w_1), (\cdot -)\} = [K(\lambda_1, \cdot)(w_1)](-) \in \mathbb{F}^{\Lambda \times W}.$$

Then for the reproducing kernel Hilbert space $\mathfrak{K}$ of the $k(\cdot\,\cdot)$ given in (C.1), the results (1) become:

$$\text{(C.3)} \quad \begin{cases} \text{(a)} \quad \mathfrak{K} \subseteq \mathbb{F}^{\Lambda \times W}, \\ \text{(b)} \quad \{T(\lambda)w: (\lambda, w) \in \Lambda \times W\} \text{ is a fundamental subset of } \mathfrak{K}, \\ \text{(c)} \quad \forall f \in \mathfrak{K} \text{ and } \forall (\lambda, w) \in \Lambda \times W, \qquad (f, T(\lambda)w)_{\mathfrak{K}} = f(\lambda, w), \\ \text{(d)} \quad \forall (\lambda_1, w_1), \qquad (\lambda_2, w_2) \in \Lambda \times W, \\ \qquad\qquad (T(\lambda_1)w_1, T(\lambda_2)w_2)_{\mathfrak{K}} = [K(\lambda_1, \lambda_2)(w_1)](w_2), \\ \text{(e)} \quad \forall (\lambda, w) \in \Lambda \times W, \qquad |T(\lambda)w|^2_{\mathfrak{K}} = [K(\lambda, \lambda)(w)](w). \end{cases}$$

We now interrupt the proof to assert the following lemma:

C.4 LEMMA. *$\forall \lambda \in \Lambda$, $T(\lambda) \in \mathrm{CL}(W, \mathfrak{K})$ and $|T(\lambda)| = \sqrt{K(\lambda, \lambda)}$. Thus $T(\cdot)$ is a function on Λ to $\mathrm{CL}(W, \mathfrak{K})$.*

Proof of C.4. Let $\lambda \in \Lambda$. Then by (C.3)(b), $\forall w \in W$, $T(\lambda)w \in \mathfrak{K}$. Hence

$$(1) \qquad T(\lambda) \text{ is a function on } W \text{ to } \mathfrak{K}.$$

It follows from (C.3)(a) that $T(\lambda)w \in \mathbb{F}^{\Lambda\times W}$; indeed from (C.2), $\forall(\lambda', w') \in \Lambda \times W$,

$$(2) \qquad [T(\lambda)(w)](\lambda', w') = [K(\lambda, \lambda')(w)](w').$$

Hence $\forall w_1, w_2 \in W$ and $\forall c_1, c_2 \in \mathbb{F}$, and we have

$$[T(\lambda)(c_1w_1 + c_2w_2)](\lambda', w') = [K(\lambda, \lambda')(c_1w_1 + c_2w_2)](w').$$

Since $K(\lambda, \lambda')$ is a linear operator on W to W^*, we easily infer that

$$[T(\lambda)(c_1w_1 + c_2w_2)](\lambda', w') = [c_1T(\lambda)w_1 + c_2T(\lambda)w_2](\lambda', w').$$

This shows of course that

$$(3) \qquad T(\lambda) \text{ is linear on } W \text{ to } \mathfrak{K}.$$

Finally, since $T(\lambda) \in \mathrm{CL}(W, \mathfrak{H})$, we have, cf. 2.2(c),

$$|T(\lambda)|^2 = |T(\lambda)^* T(\lambda)| = K(\lambda, \lambda),$$

whence

$$(4) \qquad |T(\lambda)| = \sqrt{K(\lambda, \lambda)}.$$

By (1), (3) and (4) we have the lemma. □

To resume the proof of Thm. 2.10(a), let $\lambda, \lambda' \in \Lambda$. Then by C.4, $T(\lambda), T(\lambda') \in \mathrm{CL}(W, \mathfrak{K})$. Hence, cf. 2.2(c),

$$T(\lambda')^* T(\lambda) \in \mathrm{CL}(W, W^*),$$

and $\forall w, w' \in W$,

$$\begin{aligned}[\{T(\lambda')^* T(\lambda)\}(w)](w') &= (T(\lambda)w, T(\lambda')w')_{\mathfrak{K}} \\ &= [K(\lambda, \lambda')(w)](w') \quad \text{by (C.3)(d).}\end{aligned}$$

It follows that

$$(\text{C.5}) \qquad \forall \lambda, \lambda' \in \Lambda, \qquad T(\lambda')^* T(\lambda) = K(\lambda, \lambda').$$

Now let $\mathfrak{H}$ be any Hilbert space over $\mathbb{F}$ such that $\dim. \mathfrak{H} \geqq \alpha \underset{d}{=} \dim. \mathfrak{K}$. Then $\exists$ an isometry V on $\mathfrak{K}$ to $\mathfrak{H}$. Let

$$\forall \lambda \in \Lambda, \qquad X(\lambda) \underset{d}{=} V \circ T(\lambda).$$

Then obviously $\forall \lambda, \lambda' \in \Lambda$ and $\forall w, w' \in W$,

$$(\text{C.6}) \qquad \begin{cases} X(\lambda) \in \mathrm{CL}(W, \mathfrak{H}), \\ X(\lambda')^* X(\lambda) = T(\lambda')^* V^* V T(\lambda) = T(\lambda')^* T(\lambda), \\ |X(\lambda)w|^2_{\mathfrak{H}} = |V\{T(\lambda)w\}|^2_{\mathscr{H}} = |T(\lambda)w|^2_{\mathfrak{H}}, \\ |X(\lambda)| = |T(\lambda)|. \end{cases}$$

It follows from (C.6) that the function $X(\cdot)$ is on Λ to $\mathrm{CL}(W, \mathfrak{H})$ and has all properties listed in (C.5), (C.3)(e) and C.4 for $T(\cdot)$. This proves Thm. 2.10(a) and (b). □

Appendix D. Proof of Corollary 2.12. (a) We have, writing X_λ for $X(\lambda)$,

$$\text{LHS(a)} = |X_\lambda(w)|^2 + |X_{\lambda'}(w)|^2 - (X_\lambda(w), X_{\lambda'}(w)) - (X_{\lambda'}(w), X_\lambda(w)).$$

Since by 2.2(c) and 2.8(c)

$$(X_\lambda(w), X_{\lambda'}(w)) = [(X_{\lambda'}^* X_\lambda)(w)](w) = [K(\lambda, \lambda')(w)](w),$$

we easily obtain (a).

(b) Since $X(\lambda) - X(\lambda') \in \text{CL}(W, \mathfrak{H})$, we have, cf. 2.2(c),

$$|X(\lambda) - X(\lambda')|^2 = |\{X(\lambda) - X(\lambda')\}^* \{X(\lambda) - X(\lambda')\}|.$$

But the operator-product on the RHS is, cf. 2.8(c),

$$K(\lambda, \lambda) + K(\lambda', \lambda') - K(\lambda, \lambda') - K(\lambda', \lambda).$$

Hence we have (b). □

Appendix E. Proof of Corollary 2.13. (a) Let $K(\cdot\,\cdot)$ be continuous on $\Lambda \times \Lambda$ to the Banach space $\text{Cl}(W, \mathfrak{H})$, so that $\forall(\lambda_0, \lambda_0') \in \Lambda \times \Lambda$,

$$|K(\lambda, \lambda') - K(\lambda_0, \lambda_0')|_B \to 0, \quad \text{as } (\lambda, \lambda') \to (\lambda_0, \lambda_0').$$

From this it follows easily that

$$\text{RHS } 2.12\text{(b)} \to 0, \quad \text{as } \lambda' \to \lambda.$$

Hence from 2.12(b),

$$|X(\lambda') - X(\lambda)|_B \to 0, \quad \text{as } \lambda' \to \lambda.$$

Thus (a) is proved.

(b) Let $K(\cdot\,\cdot)$ be strongly continuous on $\Lambda \times \Lambda$ to $\text{CL}(W, \mathfrak{H})$. Then $\forall(\lambda_0, \lambda_0') \in \Lambda \times \Lambda$,

$$(1) \qquad |\{K(\lambda, \lambda) - K(\lambda_0, \lambda_0')\}(w)|_{W^*} \to 0, \quad \text{as } (\lambda, \lambda') \to (\lambda_0, \lambda_0').$$

But from 2.12(a),

$$(2) \quad |X(\lambda)w - X(\lambda')w|_{\mathfrak{H}}^2 \leqq |\{K(\lambda, \lambda) + K(\lambda', \lambda') - K(\lambda, \lambda') - K(\lambda', \lambda)\}(w)|_{W^*} \cdot |w|.$$

Since by (1), RHS(2) $\to 0$, as $\lambda' \to \lambda$, we are done. □

Appendix F. Proof of Corollary 2.15. (a) Let $\mathcal{S}_X$ be the subspace of $X(\cdot)$; cf. 2.8(b). Then the linear manifold

$$(1) \qquad \mathcal{M} \underset{d}{=} \langle X(\lambda)(w) \colon (\lambda, w) \in \Lambda \times W \rangle \quad \text{is e.d. in } \mathcal{S}_X.$$

Now let $w \in W$, and grant momentarily that

$$(\text{I}) \qquad \forall(\lambda', w') \in \Lambda \times W, \qquad (X(\cdot)w, X(\lambda')(w')) \in \mathcal{M}(\mathcal{U}, \text{Bl}(\mathbb{F})).$$

Then since the last is a vector space, it follows that

$$(2) \qquad \forall x \in \mathcal{M}, \qquad (X(\cdot)w, x) \in \mathcal{M}(\mathcal{U}, \text{Bl}(\mathbb{F})).$$

But the last space is also closed under sequential pointwise convergence. Since for any $y \in \mathfrak{H}$, there is by (1) a sequence $(x_n)_1^\infty$ in $\mathcal{M}$ such that $P_{\mathcal{S}_X} y = \lim_{n\to\infty} x_n$, it follows from (2) that

$$(X(\cdot)w, y)_{\mathfrak{H}} = (X(\cdot)w, P_{\mathcal{S}_X} y)_{\mathfrak{H}} = \lim_{n\to\infty} (X(\cdot)w, x_n)_{\mathfrak{H}} \in \mathcal{M}(\mathcal{U}, \text{Bl}(\mathbb{F})),$$

as desired. Hence it only remains to prove (I).

Proof of (I). By 2.2(c) and 2.8(c)

$$(3) \qquad (X(\lambda)w, X(\lambda')w')_{\mathfrak{H}} = [\{X(\lambda')^* X(\lambda)\}(w)](w') = [K(\lambda, \lambda')(w)](w').$$

Now for $w \in W$ and any Banach space $\mathscr{X}$, let $\mathscr{E}_w^{\mathscr{X}}$ denote the restriction to $\mathrm{CL}(W, \mathscr{X})$ of the evaluation at w, i.e.

$$\forall T \in \mathrm{CL}(W, \mathscr{X}), \qquad \mathscr{E}_w^{\mathscr{X}}(T) \underset{d}{=} T(w) \in \mathscr{X}.$$

Obviously each $\mathscr{E}_w^{\mathscr{X}}$ is a continuous function on the space $\mathrm{CL}(W, \mathscr{X})$, equipped with the strong operator topology τ_s, to the Banach space $\mathscr{X}$, and therefore

$$(4) \qquad \forall w \in W, \qquad \mathscr{E}_w^{\mathscr{X}} \in \mathscr{M}(\sigma\text{-alg}(\tau_s), \mathrm{Bl}(\mathscr{X})).$$

Now note that $\mathscr{E}_w^{\mathbb{F}}$ is the restriction of the evaluation at w to $\mathrm{CL}(W, \mathbb{F})$, i.e. to W' and that the strong operator topology of $\mathrm{CL}(W, \mathbb{F})$ is precisely the so-called weak* topology τ^* for W'. Hence, with $\mathrm{Bl}(W') \underset{d}{=}$ the Borel algebra for the norm topology,

$$\forall w \in W, \quad \mathscr{E}_w^{\mathbb{F}} \in \mathscr{M}(\sigma\text{-alg}(\tau^*), \mathrm{Bl}(\mathbb{F})) \subseteq \mathscr{M}(\mathrm{Bl}(W'), \mathrm{Bl}(\mathbb{F})).$$

Letting $\forall z \in \mathbb{F}$, $C(z) \underset{d}{=} \bar{z}$, it follows easily that $\forall w' \in W$.

$$(5) \qquad \bar{\mathscr{E}}_{w'}^{\mathbb{F}} \underset{d}{=} \mathscr{E}_{w'}^{\mathbb{F}} \circ C = \text{a funct. on } W^* \text{ to } \mathbb{F} \in \mathscr{M}(\mathrm{Bl}(W^*), \mathrm{Bl}(\mathbb{F})).$$

Combining (ii), (4), (5) we see that

$$(6) \qquad \forall \lambda' \in \Lambda \text{ and } \forall w, w' \in W, \qquad \bar{\mathscr{E}}_{w'}^{\mathbb{F}} \circ \mathscr{E}_w^{W^*} \circ K^{\lambda'} \in \mathscr{M}(\mathscr{U}, \mathrm{Bl}(\mathbb{F}))$$

i.e., since by (3),

$$(X(\cdot)w, X(\lambda')w')_{\mathfrak{H}} = [K(\cdot, \lambda')(w)](w') = \bar{\mathscr{E}}_{w'}^{\mathbb{F}} \circ \mathscr{E}_w^{W^*} \circ K(\lambda'),$$

we have (I). This completes the proof of (a).

(b) It is known (cf. e.g. [23, 2.8(a)]) that $\mathscr{U}$-scalar measurability is equivalent to membership in $\mathscr{M}(\mathscr{A}, \sigma\text{-alg}(\mathscr{N}_w))$, where $\mathscr{N}_w$ is the standard base of the weak topology for $\mathfrak{H}$. Also, for separable $\mathfrak{H}$, $\sigma\text{-alg}(\mathscr{N}_w) = \mathrm{Bl}(\mathfrak{H})$; cf. e.g. [23, 2.5(b)]. It therefore follows from (a) that $\forall w \in W$, $X(\cdot)w \in \mathscr{M}(\mathscr{U}, \mathrm{BL}(\mathfrak{H}))$. □

Appendix G. Proof of Corollary 2.16. (a) Let $\lambda_1, \lambda_2 \in \Lambda$ and $c_1, c_2 \in \mathbb{F}$. Then $\forall \lambda' \in \Lambda$,

$$\begin{aligned} X(\lambda')^* \cdot X(c_1\lambda_1 + c_2\lambda_2) &= K(c_1\lambda_1 + c_2\lambda_2, \lambda') \\ &= c_1 K(\lambda_1, \lambda') + c_2 K(\lambda_2, \lambda') \\ &= c_1 X(\lambda')^* X(\lambda_1) + c_2 X(\lambda')^* \cdot X(\lambda_2) \\ &= X(\lambda')^* \{c_1 X(\lambda_1) + c_2 X(\lambda_2)\}, \end{aligned}$$

where the second equality stems from the sesquilinearity of $K(\cdot\,\cdot)$. Thus $X(\lambda')^* \cdot D = 0$, where

$$(1) \qquad D \underset{d}{=} X(c_1\lambda_1 + c_2\lambda_2) - c_1 X(\lambda_1) - c_2 X(\lambda_2) \in \mathrm{CL}(W, \mathfrak{H}).$$

Hence, cf. 2.2(c),

$$D(W) \subseteq \text{null space of } X(\lambda')^* = \{X(\lambda')(W)\}^{\perp}.$$

As this holds $\forall \lambda' \in \Lambda$, we have, in fact

$$D(W) \subseteq \bigcap_{\lambda' \in \Lambda} [\{X(\lambda')(W)\}^{\perp}] = [\mathfrak{S}\{X(\lambda')(W) \colon \lambda' \in \Lambda\}]^{\perp} = \mathscr{S}_X^{\perp}.$$

But by (1), $\forall w \in W$, $D(w) \in \mathscr{S}_X$, i.e. $D(W) \subseteq \mathscr{S}_X$. Hence $D = 0$ on W, i.e., cf. (1), $X(\cdot)$ is linear on Λ to CL$(W, \mathfrak{H})$.

(b) By 2.10(b), $\forall \lambda \in \Lambda$, $|X(\lambda)|^2 = |K(\lambda, \lambda)|$. Dividing both sides by $|\lambda|^2$, $0 \neq \lambda \in \Lambda$, and taking the sup, we see that $|X|^2 = |K| < \infty$, and therefore that the linear operator $X(\cdot)$ is continuous on Λ. □

Appendix H. Proof of Theorem 5.9. (a) Let $\forall n \in \mathbb{N}_+$, $B_n \in \mathscr{D}^{\text{loc}}$, B_n be $\|$ and $B \underset{d}{=} \bigcup_1^\infty B_n$. Then we have only to show that

$$\text{(I)} \qquad \mathscr{M}_T(B_m) \perp \mathscr{M}_T(B_n), \quad \text{for } m \neq n,$$

$$\text{(II)} \qquad \mathscr{M}_T(B) = \sum_1^\infty \mathscr{M}_T(B_n).$$

Proof of (I). Let $m \neq n$, $\Delta_1 \in \mathscr{D} \cap 2^{B_m}$, $\Delta_2 \in \mathscr{D} \cap 2^{B_n}$. Then $\Delta_1 \| \Delta_2$. Hence by (5.7), $T(\Delta_1)^* T(\Delta_2) = 0$; consequently, Range $T(\Delta_2) \subseteq$ null space $T(\Delta_1)^* = \{\text{Range } T(\Delta_1)\}^\perp$; cf. 2.2(c). This shows that

$$\forall \Delta_1 \in \mathscr{D} \cap 2^{B_m} \text{ and } \forall \Delta_2 \in \mathscr{D} \cap 2^{B_n}, \qquad T(\Delta_1)(W) \perp T(\Delta_2)(W).$$

Hence by (5.8), we have (I).

Proof of (II). Since $\forall n \in \mathbb{N}_+$, $\mathscr{D} \cap 2^{B_n} \subseteq \mathscr{D} \cap 2^B$, it follows from (5.8) that $\mathscr{M}_T(B_n) \subseteq \mathscr{M}_T(B)$, whence

$$\text{(1)} \qquad \sum_1^\infty \mathscr{M}_T(B_n) \subseteq \mathscr{M}_T(B).$$

Next, let $\Delta \in \mathscr{D} \cap 2^B$. Then

$$\Delta = \bigcup_1^\infty (\Delta \cap B_n), \quad \Delta \cap B_n \in \mathscr{D} \quad \text{and} \quad \Delta \cap B_n \|.$$

Since, cf. [21, 8.6(e)], $T(\cdot)$ is s.c.a. on $\mathscr{D}$, it follows that

$$\forall w \in W, \qquad T(\Delta)(w) = \sum_1^\infty T(\Delta \cap B_n)(w) \in \sum_1^\infty T(\Delta \cap B_n)(W).$$

Therefore

$$T(\Delta)(W) \subseteq \sum_1^\infty T(\Delta \cap B_n)(W) \subseteq \sum_1^\infty \mathscr{M}_T(B_n).$$

Now the spatial sum on the RHS is a linear manifold, and by (I) it is closed. Hence

$$\text{(2)} \qquad \mathscr{M}_T(B) \underset{d}{=} \mathfrak{S}\{T(\Delta)(W): \Delta \in \mathscr{D} \cap 2^B\} \subseteq \sum_1^\infty \mathscr{M}_T(B_n).$$

By (1) and (2), we have (II), and (a) is proved.

(b) Since $\forall B \in \mathscr{D}^{\text{loc}}$, $Q_T(B)$ is just the orthogonal projection on $\mathscr{M}_T(\Omega)$ onto $\mathscr{M}_T(B)$, (b) follows at once from (a). □

Notes added in proof.

Note 1. Thm. 2.10 is not needed before 4.15, and is stated in § 2 only for logical coherence. It is obvious that the covariance kernel $K(\cdot\,\cdot)$ of a Hilbertian variety $X(\cdot)$ on Λ to CL$(W, \mathfrak{H})$ is PD on $\Lambda \times \Lambda$ to CL(W, W^*). Thm. 2.10 is the, much deeper, converse of this result, which, as indicated in § 1, is crucial for dilation theory (4.15 et seq.).

Note 2. We state here some useful but obvious conclusions, which we missed drawing from our results in § 4.

In 4.8(d), by dividing by $|x|^{2^n}$, taking the 2^nth root and then letting $n\to\infty$, we obviously get the inequality:

$$\text{(4.8.1)}\qquad \forall x\in\langle\mathcal{D}\rangle\backslash\{0\} \text{ and } \forall p\in\Gamma_{0+},\qquad \frac{|S(p)x|}{|x|}\leqq \varliminf_{n\to\infty} |S(2^n p)x|^{2^{-n}}.$$

Combining this with 4.8(c), we also get

$$\text{(4.8.2)}\qquad \begin{cases} \forall x\in\langle\mathcal{D}\rangle\backslash\{0\} \text{ and } \forall t\in\Gamma, & \dfrac{|S(t)x|^2}{|x|^2}\leqq \varliminf_{n\to\infty} |S(2^n p)x|^{2^{-n}} \\ & \text{where } p \underset{d}{=} t^*+t. \end{cases}$$

In case $\Lambda=\Gamma$, cf. 4.12, and $X(\cdot)$ has a covariance function $R(\cdot)$, we should have noted, before turning to Thm. 4.13, that the covariance kernel $K(\cdot\cdot)$ of $X(\cdot)$ has the transfer property 4.5(a), cf. 4.12 and 4.1(b), and so the results in 4.7 and 4.8 apply; thus we have:

4.12.1 THEOREM. *Let* (i) Λ, W, $\mathfrak{H}$ *and* $X(\cdot)$ *be as in* 4.12, *and* $\mathcal{D}$ *as in* (4.6), (ii) $X(\cdot)$ *have a covariance function on* Λ *to* CL(W, W^*). *Then*

(a) $\forall t\in\Lambda$, $\exists a$ (s.v.) *closed linear operator* $S(t)$ *from* $\mathcal{S}_X$ *to* $\mathcal{S}_X$ *with domain* $\mathcal{D}_t\supseteq\langle\mathcal{D}\rangle\ni$

$\forall\lambda\in\Lambda,\qquad S(t)\cdot X(\lambda)=X(t+\lambda)$ *and* $\text{Rstr.}_{\mathcal{D}}S(t^*)\subseteq S(t)^*$;

(b) *the results* 4.8(a)–(d), (4.8.1) *and* (4.8.2) *hold for this* $S(\cdot)$.

Note 3. After this paper was submitted we learned of the paper by T. Ito, *On the commutative family of subnormal operators*, J. Fac. Sci. Hokkaido Univ., 14 (1958), pp. 1–15.

Prof. Ito considers the extensions of the Halmos and Bram theorems for semi-groups of continuous linear operators T_λ on W to W, parametrized over an abelian s.g. Λ. He works in the framework of projections rather than isometries; consequently he does not consider image-extensions (1.5 above). Apart from this one difference, however, his Thm. 1 and our Thm. 6.2 are equivalent. But whereas we deduce Thm. 6.2 directly from our strengthened version 5.3 of Nagy's Thm., Prof. Ito gets his Thm. 1 by adapting the method of Bram's original proof to his semi-group set-up.

Note 4. After this paper was submitted we received from Professor F. H. Szafraniec a reprint of his paper, *Dilations on involution semi groups*, to appear in the Proc. Amer. Math. Soc. In essence his theorem replaces our condition 4.13(β) by an interesting equivalent condition (S). We present here this theorem in our notation, along with a demonstration based on our 4.12.1 and 4.13.

THEOREM (Szafraniec). *Let* Λ, W, $\mathfrak{H}$, $X(\cdot)$ *be as in* 4.12. *Then the following conditions are equivalent*:

(α) $X(\cdot)$ *possesses a propagator* $S(\cdot)$ *on* Λ *to* CL($\mathcal{S}_X$, $\mathcal{S}_X$) *such that* $\forall t\in\Lambda$, $S(t^*)=S(t)^*$,

(β) $X(\cdot)$ *possesses a covariance function* $R(\cdot)$ *with the property that*

$$\text{(S)}\qquad \begin{cases} \exists \text{ a function } \sigma(\cdot) \text{ on } \Lambda \text{ to } \mathbb{R}_{0+} \text{ and } \exists c\in\mathbb{R}_+\ni \\ \forall s,t\in\Lambda,\qquad |R(t)|\leqq c\cdot\sigma(t) \quad\text{and}\quad \sigma(s+t)\leqq\sigma(s)\sigma(t). \end{cases}$$

Proof. (α), (β). Let (α) hold. Then $\forall t\in\Lambda$,

$$|R(t)|=|X(0)^*S(t)X(0)|\leqq|X(0)|^2\cdot|S(t)|.$$

If we let $\sigma(t) \underset{d}{=} |S(t)|$ and $c \underset{d}{=} |X(0)|^2$, it follows from the semi-group property of $S(\cdot)$ that (S) holds. Thus we have (β).

(β) ⇒ (α). Let (β) hold. Then $X(\cdot)$ has a covariance function $R(\cdot)$, and so, cf. 2.10(b),

$$(1) \quad \forall \lambda \in \Lambda \quad \text{and} \quad \forall w \in W, \qquad |X(\lambda)w|^2 = R(\lambda^* + \lambda)(w)(w) \leqq |R(\lambda^* + \lambda)| \cdot |w|^2.$$

Also by Thm. 4.12.1, the closed, densely defined, propagators $S(t)$, $t \in \Lambda$, satisfy (4.8.2), i.e. with $p \underset{d}{=} t^* + t$ we have

$$(2) \qquad \forall x \in \langle \mathcal{D} \rangle \quad \text{and} \quad \forall t \in \Lambda, \qquad |S(t)x|^2 \leqq \varliminf_{n \to \infty} |S(2^n p)x|^{2^{-n}} |x|^2.$$

Now let $a \in \Lambda$, $w \in W$ and $x \underset{d}{=} X(a)w$. Then by (1)

$$(3) \qquad |x|^2 = |X(a)w|^2 = R(a^* + a)(w)(w),$$

$$(4) \qquad |S(t)x|^2 = |X(t + a)w|^2 = R(a^* + t^* + t + a)(w)(w),$$

and

$$\begin{aligned} |S(2^n p)x|^2 = |X(2^n p + a)w|^2 &\leqq |R(a^* + 2^{n+1} p + a)| \, |w|^2 \\ &\leqq c \cdot \sigma(a^*)\sigma(a)\{\sigma(p)\}^{2^{n+1}} |w|^2, \qquad \text{by } (S). \end{aligned}$$

Taking the 2^{n+1}st root and letting $n \to \infty$, we get

$$(5) \qquad \varliminf_{n \to \infty} |S(2^n p)x|^{2^{-n}} \leqq \sigma(p).$$

Substituting from (3)–(5) in (2) yields

$$R(a^* + t + t + a)(w)(w) \leqq \sigma(t^* + t) \cdot R(a^* + a)(w)(w).$$

As this holds $\forall w \in W$, we have

$$(6) \qquad 0 \leq R(a^* + t^* + t + a) \leq \sigma(t^* + t) \cdot R(a^* + a).$$

(6) shows that with $\gamma(t) \underset{d}{=} \sigma(t + t)$, the condition 4.13(β) prevails. It follows from Thm. 4.13(a) that (α) holds. □

We see from this that Thm. 4.13 can be augmented by the addition of the condition (S) to the conditions (α), (β) in 4.13(a), and the proof just given incorporated in that of 4.13. It also seems clear that our 4.13 is provable within Professor Szafraniec's framework. Thus, in a way, Thm. 4.13 and Szafraniec's Theorem are equivalent. It should be noted, however, that for the more general case $\Lambda \neq \Gamma$, in which the involutory s.g. Γ acts on an arbitrary space Λ, there is no condition on $K(\cdot\,\cdot)$ analogous to (S) on $R(\cdot)$, and consequently Szafraniec's Thm. does not survive. It would therefore seem that our Thm. 4.10 is still the most potent result: it covers the most general case and subsumes all known theorems.

REFERENCES

[1] G. D. Allen, F. J. Narcowich and J. P. Williams, *An operator version of a theorem of Kolmogorov*, Pacific J. Math., 61 (1975), pp. 305–321.

[2] N. Aronszajn, *La theorie generale de noyaux reproduisants et ses applications*, Proc. Cambridge Philos. Soc., 39 (1944), pp. 133–153.

[3] ———, *Theory of reproducing kernels*, Trans. Amer. Math. Soc., 68 (1950), pp. 337–404.
[4] W. ARVESON, *Subalgebras of C^*-algebras*, Acta Math., 123 (1969), pp. 141–224.
[5] ———, *Subalgebras of C^*-algebras II*, Acta Math., 128 (1972), pp. 271–308.
[6] J. BRAM, *Subnormal operators*, Duke Math. J., 22 (1955), pp. 75–94.
[7] S. A. CHOBANIAN, *On a class of functions of a Banach space valued stationary stochastic process*, Sakharth SSR Mecn. Acad. Moambe, 55 (1969), pp. 21–24.
[8] ———, *On some properties of positive operator-valued measures in Banach spaces*, Ibid., 57 (1970), pp. 273–276.
[9] I. GELFAND AND D. RAIKOV, *Irreducible unitary representations of locally bicompact groups*, Math. Collection, Moscow, 13 (1943), pp. 301–316.
[10] R. K. GETOOR, *The shift operator for non-stationary stochastic processes*, Duke Math. J., 23 (1956), pp. 175–187.
[11] P. R. HALMOS, *Normal dilations and extensions of operators*, Summa Brasiliensis Math., 2 (1950), pp. 125–134.
[12] E. HILLE AND R. S. PHILLIPS, *Functional Analysis and Semi-groups*, Colloquium Publications, vol. 31, American Mathematical Society, Providence, R I, 1957.
[13] K. KARHUNEN, *Uber lineare Methoden in der Wahrscheinlichkeitsrechnung*, Ann. Acad. Sci. Fenn. Ser. A I, 37 (1947).
[14] A. Y. KHINCHINE, *Korrelationstheorie der stationaren stochastischen Prozesse*, Math. Ann., 109 (1934), pp. 604–615.
[15] A. N. KOLMOGOROV, *Stationary sequences in Hilbert space*, Bull. Math. Univ. Moscow, 2 (1941), pp. 1–40.
[16] A. LEBOW, *On von Neumann's theory of spectral sets*, J. Math. Anal. Appl., 7 (1963), pp. 64–90.
[17] J. S. MACNERNEY, *Hellinger integrals in inner product spaces*, Abstracts, Bull. Amer. Math. Soc., 61 (1955), pp. 537–539.
[18] ———, *Hellinger integrals in inner product spaces*, J. Elisha Mitchell Sci. Soc., 76 (1960), pp. 252–273.
[19] P. MASANI, *Shift invariant spaces and prediction theory*, Acta Math., 107 (1962), pp. 275–290.
[20] ———, *Orthogonally scattered measures*, Advances in Math., 2 (1968), pp. 61–117.
[21] ———, *Quasi-isometric measures and their applications*, Bull. Amer. Math. Soc., 76 (1970), pp. 427–528.
[22] ———, *On infinitely decomposable probability distributions and helical varieties in Hilbert space*, Multivariate Analysis (Proc. of the International Symposium held in Dayton, Ohio, June, 1972), P. R. Krishnaiah, ed., Academic Press, New York, 1973, pp. 209–223.
[23] ———, *Measurability and Pettis integration in Hilbert space*, Crelle's J., to appear.
[24] V. MANDREKAR AND H. SALEHI, *The square-integrability of operator-valued functions with respect to a non-negative operator-valued measure and the Kolmogorov isomorphism theorem*, Indiana Univ. Math. J., 20 (1970), pp. 545–563.
[25] A. G. MIAMEE AND H. SALEHI, *Necessary and sufficient conditions for factorability of non-negative operator-valued functions on a Banach space*, Proc. Amer. Math. Soc., 46 (1974), pp. 43–50.
[26] B. SZ.-NAGY, *Extensions of Linear Transformations in Hilbert Space Which Extend Beyond This Space*, Frederick Ungar, New York, 1960; German ed., Deutsher Verlag der Wissenchaften, Berlin, 1950.
[27] B. SZ.-NAGY AND C. FOIAS, *Harmonic Analysis of Operators on Hilbert Space*, North Holland, New York, 1970.
[28] M. A. NAIMARK, *Normed algebras*, Wolters–Noordhoff, Groningen, the Netherlands, 1972.
[29] E. PARZEN, *Regression analysis of continuous parameter time series*, Proc. Fourth Berkeley Symposium on Mathematical Statistics and Probability, 1961, vol. 1, pp. 469–489.
[30] G. B. PEDRICK, *Theory of reproducing kernels in Hilbert spaces of vector valued functions*, Univ. of Kansas Tech. Rep. 19, Lawrence, 1957.
[31] M. ROSENBERG, *Mutual subordination of multivariate stationary processes over any locally compact abelian group*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 12 (1969), pp. 333–343.
[32] W. STINESPRING, *Positive functions on C^*-algebras*, Proc. Amer. Math. Soc., 6 (1955), pp. 211–216.
[33] N. N. VAKHANIA, *The covariance of random elements in a Banach space*, Thbilis Sahelmc. Univ. Gamoqeneb. Math. Inst. Šrom., 2 (1969), pp. 179–184.
[34] J. VON NEUMANN AND I. J. SCHOENBERG, *Fourier integrals and metric geometry*, Trans. Amer. Math. Soc., 50 (1941), pp. 226–251.
[35] N. WIENER AND P. MASANI, *The prediction theory of multivariate stochastic processes, Part I: the regularity condition*, Acta Math., 98 (1957), pp. 111–150.

SIAM J. MATH. ANAL.
Vol. 9, No. 3, June 1978

Copyright © 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0903-0003$1.00/0

GENERALIZED POWER SERIES EXPANSIONS FOR A CLASS OF ORTHOGONAL POLYNOMIALS IN TWO VARIABLES*

TOM KOORNWINDER† AND IDA SPRINKHUIZEN-KUYPER†

Abstract. This paper continues the analysis of a class of orthogonal polynomials in two variables on a region bounded by two straight lines and a parabola touching these lines, which was introduced by the first author. An explicit series expansion for these polynomials is obtained, which generalizes Constantine's expansion of hypergeometric functions of (2×2) matrix argument in terms of James' zonal polynomials. In two special cases the orthogonal polynomials turn out to be Appell's hypergeometric F_4-functions and certain hypergeometric functions in two variables of order three, respectively.

1. Introduction. This paper continues the analysis of a class of orthogonal polynomials in two variables over a region bounded by two straight lines and a parabola touching these lines. The basic results on this class of polynomials are given in a paper by the first author [21], where the polynomials are introduced, and in another paper by the second author [28]. See also the survey papers [23] and [25] by the first author.

These orthogonal polynomials in two variables, which in this paper will be denoted by $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$, can be considered as highly nontrivial generalizations of the Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$. The main purpose of this paper is the derivation of an explicit series expansion for $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ which generalizes the hypergeometric power series expansion for Jacobi polynomials. Such an expansion should have the form

$$R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)=\sum_{m,l} c_{n,k;m,l}^{\alpha,\beta,\gamma} f_{m,l}^{\alpha,\beta,\gamma}(\xi,\eta), \tag{1.1}$$

where the coefficients $c_{n,k;m,l}^{\alpha,\beta,\gamma}$ and the functions $f_{m,l}^{\alpha,\beta,\gamma}(\xi,\eta)$ have to be more elementary special functions with well-known explicit expressions. It turns out that if either $k=0$ or $k=n$ the functions $f_{m,l}$ can be chosen as monomials and the coefficients then become quotients of products of gamma functions. For $k=n$ the polynomial can be identified with a terminating Appell's hypergeometric F_4-function in two variables, and for $k=0$ we obtain a certain hypergeometric function in two variables of order three.

However, if $k\neq 0$ or n then a certain choice of monomials for $f_{m,l}$ leads to rather awkward expressions for the coefficients in (1.1). In this general case the best choice for $f_{m,l}$ seems to be the so-called *James-type* zonal polynomial $Z_{m,l}^{\gamma}(\xi,\eta)$, which can be expressed in terms of Gegenbauer polynomials. Then the coefficients in (1.1) can be expressed in terms of a hypergeometric ${}_4F_3$-function of unit argument. In doing this choice we were motivated by the fact that $R_{n,n}^{\alpha,\beta,0}(\xi,\eta)$ can be identified with a hypergeometric function of (2×2) matrix argument. Constantine [10] proved that hypergeometric functions of matrix argument have a nice explicit expansion in terms of the zonal polynomials introduced by James [18]. In the (2×2) case these zonal polynomials can be identified with our polynomials $Z_{m,l}^{0}(\xi,\eta)$.

We already pointed out that the polynomials $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ become more simple on the boundary lines $n=0$ and $n=k$ of the region $\{(n,k)\in\mathbb{Z}^2 \mid n\geqq k\geqq 0\}$ for which $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ is defined. Similarly, the analysis of the polynomials on the boundary of the orthogonality region in the (ξ,η) plane is easier than in the interior of this region. In particular, $R_{n,n}^{\alpha,\beta,\gamma}(\xi,0)$ and $R_{n,0}^{\alpha,\beta,\gamma}(\xi,\frac{1}{4}\xi^2)$ turn out to be Jacobi polynomials of argument $1-2\xi$ and $1-\xi$, respectively. Our proofs exploit these degeneracies in the

* Received by the editors July 21, 1976.
† Mathematisch Centrum, Amsterdam, the Netherlands.

(n, k) and (ξ, η) planes. The two pairs of differential recurrence relations derived in [21] and [28] will also be used as essential tools.

The results in this paper are not only an interesting part of the analysis of the polynomials $R_{n,k}^{\alpha,\beta,\gamma}(\xi, \eta)$, but they may be important for a wider class of readers. First, there is a close relationship with the theory of Jacobi polynomials. Many results for Jacobi polynomials will be used in this paper, and, on the other hand, some known results for Jacobi polynomials can be better understood from our two-variable point of view. Second, we bring some unity in the bewildering variety of special functions in more than one variable by identifying hypergeometric functions in two variables (in particular F_4) and hypergeometric functions of (2×2) matrix argument with special cases of our polynomials.

In §§ 2 and 3 of this paper we summarize the results on Jacobi polynomials and on the polynomials $R_{n,k}^{\alpha,\beta,\gamma}(\xi, \eta)$ which will be needed. In § 4 the James-type zonal polynomials are introduced. The boundary values of the polynomials $R_{n,k}^{\alpha,\beta,\gamma}(\xi, \eta)$ are considered in § 5. Section 6 contains the expansion of the polynomials $R_{n,k}^{\alpha,\beta,\gamma}(\xi, \eta)$ in terms of James-type zonal polynomials. In § 7 we consider expansions of the form (1.1) with another natural choice for the functions $f_{m,l}^{\alpha,\beta,\gamma}(\xi, \eta)$. This leads to the identification of $R_{n,n}^{\alpha,\beta,\gamma}(\xi, \eta)$ with Appell's F_4-function. Finally, in § 8 we derive expansions of $R_{n,k}^{\alpha,\beta,\gamma}(\xi, \eta)$, $k=0$ or n, as double Jacobi series with positive coefficients. Sections 7 and 8 can be read independently of each other and of § 4 and 6. However, § 5 is needed for all the subsequent sections.

In a forthcoming paper we will extend the correspondence between $R_{n,n}^{\alpha,\beta,\gamma}(\xi, \eta)$ and Appell's function F_4 to the nonpolynomial case. An expansion in terms of James-type zonal polynomials will be given for those solutions of the system of partial differential equations for F_4 which are regular in the singular point $(1, 0)$. For special values of the parameters these second solutions are precisely the hypergeometric functions of (2×2) matrix argument.

2. Properties of Jacobi polynomials. In this section we collect all results on Jacobi polynomials which will be needed in this paper. The standard formulas for Jacobi polynomials have been taken from Szegö [29, Chap. 4] and Erdélyi [13, Chap. 10]. A useful survey of many recent results on Jacobi polynomials is given by Askey [2].

Let $\alpha, \beta > -1$. Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ are orthogonal polynomials on the interval $(-1, 1)$ with respect to the weight function $(1-x)^\alpha(1+x)^\beta$ and with the normalization $P_n^{(\alpha,\beta)}(1) := (\alpha+1)_n/n!$ We will use the renormalized Jacobi polynomials $R_n^{(\alpha,\beta)}(x) := P_n^{(\alpha,\beta)}(x)/P_n^{(\alpha,\beta)}(1)$.

Differentiation formulas:

$$(1-x^2)\frac{d^2}{dx^2}R_n^{(\alpha,\beta)}(x)+(\beta-\alpha-(\alpha+\beta+2)x)\frac{d}{dx}R_n^{(\alpha,\beta)}(x) + n(n+\alpha+\beta+1)R_n^{(\alpha,\beta)}(x)=0, \tag{2.1}$$

$$\frac{d}{dx}R_n^{(\alpha,\beta)}(x)=\begin{cases}\dfrac{n(n+\alpha+\beta+1)}{2(\alpha+1)}R_{n-1}^{(\alpha+1,\beta+1)}(x) & \text{if } n>0,\\ 0 & \text{if } n=0,\end{cases} \tag{2.2}$$

$$(1-x)^{-\alpha}(1+x)^{-\beta}\frac{d}{dx}[(1-x)^{\alpha+1}(1+x)^{\beta+1}R_{n-1}^{(\alpha+1,\beta+1)}(x)] = -2(\alpha+1)R_n^{(\alpha,\beta)}(x). \tag{2.3}$$

Series expansions:

$$R_n^{(\alpha,\beta)}(x) = {}_2F_1(-n, n+\alpha+\beta+1; \alpha+1; \tfrac{1}{2}(1-x)) = \sum_{k=0}^{n} \frac{(-n)_k(n+\alpha+\beta+1)_k}{(\alpha+1)_k k!}\left(\frac{1-x}{2}\right)^k, \tag{2.4}$$

$$R_n^{(\alpha,\beta)}(x) = \left(\frac{1+x}{2}\right)^n {}_2F_1\left(-n, -n-\beta; \alpha+1; \frac{x-1}{x+1}\right) = 2^{-n} \sum_{k=0}^{n} \frac{(-n)_k(-n-\beta)_k}{(\alpha+1)_k k!}(x+1)^{n-k}(x-1)^k, \tag{2.5}$$

$$\left(\frac{1-x}{2}\right)^n = \frac{(\alpha+1)_n}{(\alpha+\beta+2)_n} \sum_{k=0}^{n} \frac{(2k+\alpha+\beta+1)(-n)_k(\alpha+\beta+2)_k}{(k+\alpha+\beta+1)(n+\alpha+\beta+2)_k k!} R_k^{(\alpha,\beta)}(x). \tag{2.6}$$

Value for $x = -1$:

$$R_n^{(\alpha,\beta)}(-1) = (-1)^n \frac{(\beta+1)_n}{(\alpha+1)_n}. \tag{2.7}$$

Linear and quadratic transformations:

$$\frac{R_n^{(\alpha,\beta)}(-x)}{R_n^{(\alpha,\beta)}(-1)} = R_n^{(\beta,\alpha)}(x), \tag{2.8}$$

$$R_{2n}^{(\alpha,\alpha)}(x) = R_n^{(\alpha,-1/2)}(2x^2-1), \tag{2.9}$$

$$R_{2n+1}^{(\alpha,\alpha)}(x) = xR_n^{(\alpha,1/2)}(2x^2-1). \tag{2.10}$$

Gegenbauer and Chebyshev polynomials:

$$R_n^{(\gamma,\gamma)}(x) = \frac{(\gamma+\frac{1}{2})_n}{(2\gamma+1)_n} \sum_{k=0}^{[n/2]} \frac{(-n)_{2k}}{(-n-\gamma+\frac{1}{2})_k k!}(2x)^{n-2k}, \tag{2.11}$$

$$R_n^{(-1/2,-1/2)}(\cos\theta) = \cos n\theta, \tag{2.12}$$

$$R_n^{(1/2,1/2)}(\cos\theta) = \frac{\sin[(n+1)\theta]}{(n+1)\sin\theta}. \tag{2.13}$$

It follows from these last two formulas that

$$R_n^{(-1/2,-1/2)}(\tfrac{1}{2}(t+t^{-1})) = \tfrac{1}{2}(t^n+t^{-n}), \tag{2.14}$$

$$R_n^{(1/2,1/2)}(\tfrac{1}{2}(t+t^{-1})) = \frac{t^{n+1}-t^{-n-1}}{(n+1)(t-t^{-1})}. \tag{2.15}$$

Formula (2.12) is a special case of

$$R_n^{(\gamma,\gamma)}(\cos\theta) = \frac{n!}{(2\gamma+1)_n} \sum_{k=0}^{n} \frac{(\gamma+\frac{1}{2})_k(\gamma+\frac{1}{2})_{n-k}}{k!(n-k)!} \cos[(n-2k)\theta], \tag{2.16}$$

which formula results in

$$R_n^{(\gamma,\gamma)}(\tfrac{1}{2}(t+t^{-1})) = \frac{n!}{(2\gamma+1)_n} \sum_{k=0}^{n} \frac{(\gamma+\frac{1}{2})_k(\gamma+\frac{1}{2})_{n-k}}{k!(n-k)!} t^{n-2k}. \tag{2.17}$$

Quadratic norm: Let

$$\omega_n^{(\alpha,\beta)} := \frac{\int_{-1}^{1}(1-x)^\alpha(1+x)^\beta\,dx}{\int_{-1}^{1}(R_n^{(\alpha,\beta)}(x))^2(1-x)^\alpha(1+x)^\beta\,dx}. \tag{2.18}$$

Then

$$\omega_n^{(\alpha,\beta)} = \frac{(2n+\alpha+\beta+1)(\alpha+1)_n(\alpha+\beta+2)_n}{(n+\alpha+\beta+1)(\beta+1)_n n!}. \tag{2.19}$$

Christoffel-Darboux formula:

$$\sum_{k=0}^{n}\omega_k^{(\alpha,\beta)}R_k^{(\alpha,\beta)}(x)R_k^{(\alpha,\beta)}(y) = \frac{2(\alpha+1)(\alpha+2)_n(\alpha+\beta+2)_n}{(2n+\alpha+\beta+2)(\beta+1)_n n!} \cdot \frac{R_{n+1}^{(\alpha,\beta)}(x)R_n^{(\alpha,\beta)}(y)-R_n^{(\alpha,\beta)}(x)R_{n+1}^{(\alpha,\beta)}(y)}{x-y}. \tag{2.20}$$

Limit formulas:

$$\lim_{\beta\to\infty}\frac{R_n^{(\alpha,\beta)}(x)}{R_n^{(\alpha,\beta)}(-1)} = \left(\frac{1-x}{2}\right)^n, \tag{2.21}$$

$$\lim_{\alpha\to\infty}R_n^{(\alpha,\beta)}(x) = \left(\frac{1+x}{2}\right)^n, \tag{2.22}$$

$$\lim_{\gamma\to\infty}R_n^{(\gamma,\gamma)}(x) = x^n. \tag{2.23}$$

These three results follow from (2.4), (2.7), (2.8) and (2.11).

Another pair of differential recurrence relations:

$$(1-x)^{1-\alpha}\frac{d}{dx}[(1-x)^\alpha R_n^{(\alpha,\beta)}(x)] = -\alpha R_n^{(\alpha-1,\beta+1)}(x), \tag{2.24}$$

$$(1+x)^{-\beta}\frac{d}{dx}[(1+x)^{\beta+1}R_n^{(\alpha-1,\beta+1)}(x)] = \alpha^{-1}(n+\alpha)(n+\beta+1)R_n^{(\alpha,\beta)}(x). \tag{2.25}$$

It follows from Slater [27, (2.5.31)] that

$$R_n^{(\alpha,\beta)}(x)R_n^{(\beta,\alpha)}(x) = {}_4F_3\left(\begin{matrix}-n, \tfrac{1}{2}(\alpha+\beta+1), \tfrac{1}{2}(\alpha+\beta+2), n+\alpha+\beta+1; \\ \alpha+1, \beta+1, \alpha+\beta+1;\end{matrix}\; 1-x^2\right). \tag{2.26}$$

In particular:

$$(R_n^{(\alpha,\alpha)}(x))^2 = {}_3F_2(-n, \alpha+\tfrac{1}{2}, n+2\alpha+1; \alpha+1, 2\alpha+1; 1-x^2). \tag{2.27}$$

Appell's hypergeometric function F_4 is defined by

$$F_4(a,b;c,c';x,y) := \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(a)_{m+n}(b)_{m+n}}{(c)_m(c')_n m!\,n!}x^m y^n, \qquad |x|^{1/2}+|y|^{1/2}<1. \tag{2.28}$$

A result of Watson (cf. Slater [27, (8.4.4)], gives

$$R_n^{(\alpha,\beta)}(x)R_n^{(\beta,\alpha)}(y) = F_4(-n, n+\alpha+\beta+1; \alpha+1, \beta+1; \tfrac{1}{4}(1-x)(1+y), \tfrac{1}{4}(1+x)(1-y)). \tag{2.29}$$

THEOREM 2.1 (cf. Bateman [4, pp. 392, 393] and Koornwinder [22, § 2]). *If*

$$R_n^{(\alpha,\beta)}(x)=\sum_{k=0}^{n} c_k\left(\frac{1+x}{2}\right)^k \tag{2.30}$$

then

$$R_n^{(\alpha,\beta)}(x)R_n^{(\alpha,\beta)}(y)=\sum_{k=0}^{n} c_k\left(\frac{x+y}{2}\right)^k R_k^{(\alpha,\beta)}\left(\frac{1+xy}{x+y}\right) \tag{2.31}$$

and

$$c_k=\frac{(-1)^n(\beta+1)_n(-n)_k(n+\alpha+\beta+1)_k}{(\alpha+1)_n(\beta+1)_k k!}. \tag{2.32}$$

The product formula and addition formula for Gegenbauer polynomials (cf. Erdélyi [12, 3.15(19) and 3.15(20)]):

$$\begin{aligned} R_n^{(\gamma,\gamma)}(x)R_n^{(\gamma,\gamma)}(y)&=\frac{\Gamma(\gamma+1)}{\pi^{1/2}\Gamma(\gamma+\frac{1}{2})}\\ &\quad\cdot\int_{-1}^{1} R_n^{(\gamma,\gamma)}(xy+(1-x^2)^{1/2}(1-y^2)^{1/2}t)(1-t^2)^{\gamma-1/2}\,dt, \end{aligned} \tag{2.33}$$

$$\gamma>-\tfrac{1}{2},$$

$$\begin{aligned} &R_n^{(\gamma,\gamma)}(xy+(1-x^2)^{1/2}(1-y^2)^{1/2}t)\\ &\quad=\sum_{k=0}^{n}\frac{(-1)^k(-n)_k(n+2\gamma+1)_k}{2^{2k}(\gamma+1)_k(\gamma+1)_k}\\ &\quad\cdot(1-x^2)^{k/2}R_{n-k}^{(\gamma+k,\gamma+k)}(x)(1-y^2)^{k/2}R_{n-k}^{(\gamma+k,\gamma+k)}(y)\\ &\quad\cdot\omega_k^{(\gamma-1/2,\gamma-1/2)}R_k^{(\gamma-1/2,\gamma-1/2)}(t). \end{aligned} \tag{2.34}$$

The product formula for Jacobi polynomials (cf. Koornwinder [22, (3.7)]):

$$\begin{aligned} R_n^{(\alpha,\beta)}(x)R_n^{(\alpha,\beta)}(y)&=\frac{2\Gamma(\alpha+1)}{\pi^{1/2}\Gamma(\alpha-\beta)\Gamma(\beta+\frac{1}{2})}\\ &\quad\cdot\int_0^1\int_0^\pi R_n^{(\alpha,\beta)}(\tfrac{1}{2}(1+x)(1+y)+\tfrac{1}{2}(1-x)(1-y)r^2\\ &\qquad+(1-x^2)^{1/2}(1-y^2)^{1/2}r\cos\phi-1)\\ &\qquad\cdot(1-r^2)^{\alpha-\beta-1}r^{2\beta+1}(\sin\phi)^{2\beta}\,dr\,d\phi,\qquad \alpha>\beta>-\tfrac{1}{2}. \end{aligned} \tag{2.35}$$

THEOREM 2.2 (cf. Szegö [29, Thm. 7.32.1]).
(a) *If* $\alpha\geqq\beta$ *and* $\alpha\geqq-\frac{1}{2}$ *then*

$$|R_n^{(\alpha,\beta)}(x)|\leqq 1\quad \textit{for } -1\leqq x\leqq 1.$$

(b) *If* $\alpha\leqq\beta$ *and* $\beta\geqq-\frac{1}{2}$ *then*

$$|R_n^{(\alpha,\beta)}(x)|\leqq|R_n^{(\alpha,\beta)}(-1)|\quad \textit{for } -1\leqq x\leqq 1.$$

The coefficients in

$$R_n^{(\alpha,\beta)}(x)=\sum_{k=0}^{n} g_{n;k}^{\alpha,\beta;a,b}R_k^{(a,b)}(x) \tag{2.36}$$

are called *connection coefficients*. We have

$$g_{n;k}^{\alpha,\beta;a,\beta}=\frac{n!(\beta+1)_n(\alpha-a)_{n-k}(n+\alpha+\beta+1)_k\omega_k^{(a,\beta)}}{(\alpha+1)_n(a+\beta+2)_n(n-k)!(n+a+\beta+2)_k}; \tag{2.37}$$

cf. Szegö [29, (9.41)]. Hence

$$g_{n;k}^{\alpha,\beta;a,\beta}>0 \quad \text{if } \alpha>a. \tag{2.38}$$

In the general case the connection coefficients are given by

$$\begin{aligned} g_{n;k}^{\alpha,\beta;a,b}&=\frac{(n+\alpha+\beta+1)_k(a+1)_k n!}{(k+a+b+1)_k(\alpha+1)_k(n-k)!k!}\\ &\cdot {}_3F_2\left(\begin{matrix}-n+k,\, n+k+\alpha+\beta+1,\, k+a+1;\\ 2k+a+b+2,\, k+\alpha+1;\end{matrix}\ 1\right); \end{aligned} \tag{2.39}$$

cf. Feldheim [14] or Askey and Gasper [3, (2.5), (2.6)].

THEOREM 2.3. *If $a\leqq b$, $\alpha+\beta\geqq a+b$, and $\beta-\alpha\leqq b-a$ then $g_{n;k}^{\alpha,\beta;a,b}\geqq 0$ and the inequality is strict except if $a=b$, $\alpha=\beta$ and $n-k$ is odd.*

This theorem is essentially a part of Theorem 1 in Askey and Gasper [3]. The last statement in Theorem 2.3 is a slight refinement of their result. It follows immediately from the recurrence relation (2.2) in [3].

The coefficients in

$$R_m^{(\alpha,\beta)}(x)R_n^{(\alpha,\beta)}(x)=\sum_{k=|m-n|}^{m+n}A_{m,n,k}^{(\alpha,\beta)}\omega_k^{(\alpha,\beta)}R_k^{(\alpha,\beta)}(x) \tag{2.40}$$

are called *linearization coefficients*. We have $A_{m,n,k}^{(\alpha,\alpha)}=0$ if $m+n+k$ is odd and

$$\begin{aligned} A_{m,n,k}^{(\alpha,\alpha)}&=\frac{(2\alpha+1)_{(m+n+k)/2}}{(\alpha+\frac{3}{2})_{(m+n+k)/2}}\\ &\cdot\frac{(\alpha+\frac12)_{(m+n-k)/2}(\alpha+\frac12)_{(n+k-m)/2}(\alpha+\frac12)_{(k+m-n)/2}m!n!k!}{(\frac12(m+n-k))!(\frac12(n+k-m))!(\frac12(k+m-n))!(2\alpha+1)_m(2\alpha+1)_n(2\alpha+1)_k} \end{aligned} \tag{2.41}$$

if $m+n+k$ is even. Formula (2.41) was first stated by Dougall [11] without proof. See Askey [2, Lecture 5] for a survey of several proofs of (2.41) which were published afterwards.

THEOREM 2.4 (cf. Gasper [15]). *If $\alpha\geqq\beta$ and $\alpha+\beta\geqq -1$ then $A_{m,n,k}^{(\alpha,\beta)}\geqq 0$.*

In [16] Gasper extended this nonnegativity result for the linearization coefficients to a slightly larger region of the (α,β) plane.

LEMMA 2.5. *Let the polynomials $p_n(x)$, $n=0,1,2,\cdots$, be orthogonal on the interval (a,b) with respect to the strictly positive weight function $w(x)$. Then any polynomial of the form*

$$f(x):=\sum_{m=k}^{n}c_m p_m(x),\quad \textit{with } c_n\neq 0,$$

has at least k zeros of odd multiplicity on (a,b).

Proof. Suppose that $f(x)$ has only l zeros $x_1,\cdots,x_l$ of odd multiplicity on (a,b) with $l<k$. Then $f(x)(x-x_1)\cdots(x-x_l)$ is either nonnegative or nonpositive on (a,b) and not identically zero. But

$$\int_a^b f(x)(x-x_1)\cdots(x-x_l)w(x)\,dx=0.$$

This is a contradiction. □

Finally we mention Saalschütz's formula (cf. Slater [27, § 2.3.1]):

(2.42) $$ {}_3F_2(a, b, -n; c, 1+a+b-c-n; 1) = \frac{(c-a)_n(c-b)_n}{(c)_n(c-a-b)_n}, \qquad n = 0, 1, 2, \cdots. $$

3. Some earlier results for the orthogonal polynomials $R_{n,k}^{\alpha,\beta,\gamma}(\xi, \eta)$. The purpose of this section is to summarize some of the results on the polynomials $p_{n,k}^{\alpha,\beta,\gamma}(u, v)$ obtained in Koornwinder [21] and Sprinkhuizen [28]. We will change the notation used in these two papers by introducing new coordinates $\xi := 1-\frac{1}{2}u$, $\eta := \frac{1}{4}(1-u+v)$ and by renormalizing the polynomials such that they are equal to 1 in the vertex $(\xi, \eta) = (0, 0)$. A motivation of this new notation will be given in § 4.

Let Ω be the region

(3.1) $$ \Omega := \{(\xi, \eta) | \eta > 0, 1-\xi+\eta > 0, \xi^2 - 4\eta > 0, 0 < \xi < 2\}, $$

which is bounded by two straight lines and a parabola touching these lines (cf. Fig. 1).

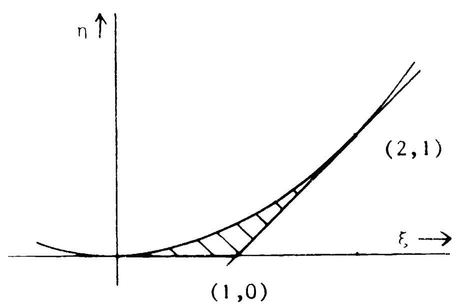

FIG. 1

Let

(3.2) $$ w_{\alpha,\beta,\gamma}(\xi, \eta) := \eta^\alpha(1-\xi+\eta)^\beta(\xi^2-4\eta)^\gamma, (\xi, \eta) \in \Omega. $$

DEFINITION 3.1. Let $\alpha, \beta, \gamma > -1$, $\alpha+\gamma+\frac{3}{2} > 0$, $\beta+\gamma+\frac{3}{2} > 0$. Let n, k be integers, $n \geqq k \geqq 0$. Then $R_{n,k}^{\alpha,\beta,\gamma}(\xi, \eta)$ is a linear combination of monomials $1, \xi, \eta, \xi^2, \xi\eta, \eta^2, \xi^3, \xi^2\eta, \cdots, \xi^n, \xi^{n-1}\eta, \cdots, \xi^{n-k}\eta^k$ such that

(i) $$ \iint_\Omega R_{n,k}^{\alpha,\beta,\gamma}(\xi, \eta)\xi^{m-l}\eta^l w_{\alpha,\beta,\gamma}(\xi, \eta)\, d\xi\, d\eta = 0 $$

if $m \geqq l \geqq 0$ and if either $m < n$ or $m = n$, $l < k$;

(ii) $$ R_{n,k}^{\alpha,\beta,\gamma}(0, 0) = 1. $$

If $p_{n,k}^{\alpha,\beta,\gamma}(u, v)$ is defined as in [21] then

(3.3) $$ R_{n,k}^{\alpha,\beta,\gamma}(\xi, \eta) = \frac{p_{n,k}^{\alpha,\beta,\gamma}(2-2\xi, 1-2\xi+4\eta)}{p_{n,k}^{\alpha,\beta,\gamma}(2, 1)}, $$

where the value of $p_{n,k}^{\alpha,\beta,\gamma}(2, 1)$ is given in [28, (7.3)].

For $\gamma = \pm\frac{1}{2}$ the polynomials $R_{n,k}^{\alpha,\beta,\gamma}(\xi, \eta)$ can be expressed in terms of Jacobi polynomials by

(3.4) $$ R_{n,k}^{\alpha,\beta,-1/2}(x+y, xy) = \tfrac{1}{2}\{R_n^{(\alpha,\beta)}(1-2x)R_k^{(\alpha,\beta)}(1-2y) + R_k^{(\alpha,\beta)}(1-2x)R_n^{(\alpha,\beta)}(1-2y)\}, $$

(3.5)
$$\begin{aligned}R_{n,k}^{\alpha,\beta,1/2}(x+y,xy)=&\frac{-(\alpha+1)}{(n-k+1)(n+k+\alpha+\beta+2)(x-y)}\\&\cdot\{R_{n+1}^{(\alpha,\beta)}(1-2x)R_k^{(\alpha,\beta)}(1-2y)\\&-R_k^{(\alpha,\beta)}(1-2x)R_{n+1}^{(\alpha,\beta)}(1-2y)\}.\end{aligned}$$

By comparing (2.20) with (3.5) we can conclude that

(3.6)
$$\begin{aligned}&R_{n,n}^{\alpha,\beta,1/2}(1-\tfrac12(x+y),\tfrac14(1-x)(1-y))\\&\qquad=\frac{(\beta+1)_n n!}{(\alpha+2)_n(\alpha+\beta+2)_n}\sum_{k=0}^n\omega_k^{(\alpha,\beta)}R_k^{(\alpha,\beta)}(x)R_k^{(\alpha,\beta)}(y).\end{aligned}$$

Let $\partial_{x_1x_2\cdots x_k}$ denote the partial derivative $\partial^k/\partial x_1\,\partial x_2\cdots\partial x_k$. Consider the second order differential operators

(3.7) $$D_-^\gamma:=\tfrac14\{\partial_{\xi\xi}+\xi\,\partial_{\xi\eta}+\eta\,\partial_{\eta\eta}+(\gamma+\tfrac32)\,\partial_\eta\},$$

(3.8) $$D_+^{\alpha,\beta,\gamma}:=16(w_{\alpha,\beta,\gamma}(\xi,\eta))^{-1}D_-^{-\gamma}\circ w_{\alpha+1,\beta+1,\gamma}(\xi,\eta),$$

(3.9) $$E_-^{\alpha,\beta}:=\tfrac12\{(1-\xi)\,\partial_{\xi\xi}-2\eta\,\partial_{\xi\eta}-\eta\,\partial_{\eta\eta}-(\alpha+\beta+2)\,\partial_\xi-(\alpha+1)\,\partial_\eta\},$$

(3.10) $$E_+^{\alpha,\beta,\gamma}:=4(w_{\alpha,\beta,\gamma}(\xi,\eta))^{-1}E_-^{-\alpha,-\beta}\circ w_{\alpha,\beta,\gamma+1}(\xi,\eta).$$

The operators $D_+^{\alpha,\beta,\gamma}$ and $E_+^{\alpha,\beta,\gamma}$ can be written more explicitly as

(3.11)
$$\begin{aligned}D_+^{\alpha,\beta,\gamma}=&16\eta(1-\xi+\eta)D_-^\gamma+4\{(\alpha+1)\xi(1-\xi)+(\alpha+\beta+2)\xi\eta-2(\beta+1)\eta\}\,\partial_\xi\\&+4\eta\{-(2\alpha+\beta+3)\xi+2(\alpha+\beta+2)\eta+2(\alpha+1)\}\,\partial_\eta\\&-4(\alpha+1)(\alpha+\beta+\gamma+\tfrac52)\xi+4(\alpha+\beta+2)(\alpha+\beta+\gamma+\tfrac52)\eta\\&+4(\alpha+1)(\alpha+\gamma+\tfrac32),\end{aligned}$$

(3.12)
$$\begin{aligned}E_+^{\alpha,\beta,\gamma}=&4(\xi^2-4\eta)\{E_-^{\alpha,\beta}-(\gamma+1)(2\partial_\xi+\partial_\eta)\}+4(\gamma+1)(\xi-2\eta)(2\partial_\xi+\xi\partial_\eta)\\&-4(\alpha+\beta+2\gamma+3)(\gamma+1)\xi+8(\gamma+1)(\alpha+\gamma+\tfrac32).\end{aligned}$$

It was proved in [21, § 5] and [28, § 4] that these differential operators act on $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ as follows:

(3.13)
$$\begin{aligned}&D_-^\gamma R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta)=0,\\&D_-^\gamma R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)=\frac{k(k+\alpha+\beta+1)(n+\gamma+\frac12)(n+\alpha+\beta+\gamma+\frac32)}{4(\alpha+1)(\alpha+\gamma+\frac32)}R_{n-1,k-1}^{\alpha+1,\beta+1,\gamma}(\xi,\eta)\\&\hspace{30em}\text{if } k>0,\end{aligned}$$

(3.14) $$D_+^{\alpha,\beta,\gamma}R_{n-1,k-1}^{\alpha+1,\beta+1,\gamma}(\xi,\eta)=4(\alpha+1)(\alpha+\gamma+\tfrac32)R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta),\qquad k>0,$$

(3.15)
$$\begin{aligned}&E_-^{\alpha,\beta}R_{n,n}^{\alpha,\beta,\gamma}(\xi,\eta)=0,\\&E_-^{\alpha,\beta}R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)\\&\quad=\frac{(n-k)(n-k+2\gamma+1)(n+k+\alpha+\beta+1)(n+k+\alpha+\beta+2\gamma+2)}{8(\gamma+1)(\alpha+\gamma+\frac32)}\\&\qquad\cdot R_{n-1,k}^{\alpha,\beta,\gamma+1}(\xi,\eta)\hspace{20em}\text{if } n>k,\end{aligned}$$

(3.16) $$E_+^{\alpha,\beta,\gamma}R_{n-1,k}^{\alpha,\beta,\gamma+1}(\xi,\eta)=8(\gamma+1)(\alpha+\gamma+\tfrac32)R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta),\qquad n>k.$$

For the calculation of the coefficients in these four differentiation formulas we used the explicit value of $p_{n,k}^{\alpha,\beta,\gamma}(2,1)$; cf. [28, (7.3)]. Note that these coefficients are nonzero if α, β, γ satisfy the inequalities of Definition 3.1.

THEOREM 3.2 (cf. [28, Thm. 8.1]). *In the power series*

$$R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)=\sum_{m,l} c_{m,l}\xi^{m-l}\eta^{l}$$

the coefficient $c_{m,l}$ is nonzero only if $m \leqq n$ and $m+l \leqq n+k$ (cf. Fig. 2).

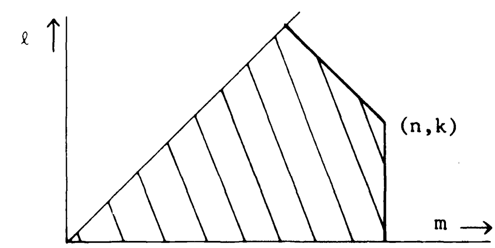

FIG. 2

This theorem also follows from the results of § 4; cf. Remark 4.4.

Analogous to (2.7) and (2.8) we have

$$\text{(3.17)}\qquad R_{n,k}^{\alpha,\beta,\gamma}(2,1)=\frac{(-1)^{n-k}(\beta+1)_k(\beta+\gamma+\frac{3}{2})_n}{(\alpha+1)_k(\alpha+\gamma+\frac{3}{2})_n},$$

$$\text{(3.18)}\qquad \frac{R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)}{R_{n,k}^{\alpha,\beta,\gamma}(2,1)}=R_{n,k}^{\beta,\alpha,\gamma}(2-\xi,1-\xi+\eta).$$

Note that the mapping $(\xi,\eta)\to(2-\xi,1-\xi+\eta)$ is a nonorthogonal reflection which maps Ω onto itself, $(0,0)$ to $(2,1)$ and which leaves the points of the line $\xi=1$ invariant.

Finally we mention the quadratic transformation formulas

$$\text{(3.19)}\qquad R_{n+k,n-k}^{\alpha,\alpha,\gamma}(\xi,\eta)=R_{n,k}^{\gamma,-1/2,\alpha}(2\xi-4\eta,\xi^2-4\eta),$$

$$\text{(3.20)}\qquad R_{n+k+1,n-k}^{\alpha,\alpha,\gamma}(\xi,\eta)=(1-\xi)R_{n,k}^{\gamma,1/2,\alpha}(2\xi-4\eta,\xi^2-4\eta).$$

The quadratic transformation $(\xi,\eta)\to(2\xi-4\eta,\xi^2-4\eta)$ maps both connected components of $\{(\xi,\eta)\in\Omega \mid \xi\neq 1\}$ onto Ω. In fact, $(0,0)$ and $(2,1)$ are both mapped to $(0,0)$, $(1,0)$ is mapped to $(2,1)$, and $(1,\frac{1}{4})$ is mapped to $(1,0)$.

Note that formulas (3.19), (3.20) and (3.4) together imply that

$$\text{(3.21)}\qquad R_{n,k}^{-1/2,-1/2,\gamma}(1-xy,\tfrac{1}{4}(x-y)^2)=\tfrac{1}{2}\{R_{n+k}^{(\gamma,\gamma)}(x)R_{n-k}^{(\gamma,\gamma)}(y)+R_{n-k}^{(\gamma,\gamma)}(x)R_{n+k}^{(\gamma,\gamma)}(y)\},$$

$$\text{(3.22)}\qquad \begin{aligned}&R_{n,k}^{-1/2,1/2,\gamma}(1-xy,\tfrac{1}{4}(x-y)^2)\\&=(x+y)^{-1}\{R_{n+k+1}^{(\gamma,\gamma)}(x)R_{n-k}^{(\gamma,\gamma)}(y)+R_{n-k}^{(\gamma,\gamma)}(x)R_{n+k+1}^{(\gamma,\gamma)}(y)\}.\end{aligned}$$

4. James-type zonal polynomials. As was pointed out in § 1, the main problem to be solved in this paper is the derivation of an explicit expression of $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ in terms of certain polynomials $Z_{m,l}^{\gamma}(\xi,\eta)$ called James-type zonal polynomials. In this section we introduce these polynomials $Z_{m,l}^{\gamma}(\xi,\eta)$, we give some motivation for the choice of these polynomials and we derive some simple properties of the expansion coefficients. To a large extent, the contents of this section coincide with Koornwinder [25, § 4.4].

DEFINITION 4.1. Let $\gamma > -1$. Let n, k be integers such that $n \geqq k \geqq 0$. Then the *James-type zonal polynomial* $Z_{n,k}^{\gamma}(\xi, \eta)$ is defined by

$$Z_{n,k}^{\gamma}(\xi, \eta) := \frac{(2\gamma+1)_{n-k}}{(\gamma+\frac{1}{2})_{n-k}} \eta^{(n+k)/2} R_{n-k}^{(\gamma,\gamma)}(\tfrac{1}{2}\eta^{-1/2}\xi). \tag{4.1}$$

It follows from (2.11) that

$$Z_{n,k}^{\gamma}(\xi, \eta) = \sum_{i=0}^{[(n-k)/2]} \frac{(-n+k)_{2i}}{(-n+k-\gamma+\frac{1}{2})_i i!} \xi^{n-k-2i} \eta^{k+i}. \tag{4.2}$$

Note that

$$Z_{n,k}^{\gamma}(\xi, \eta) = \xi^{n-k}\eta^k + \text{polynomial of degree less than } n, \tag{4.3}$$

$$\lim_{\gamma\to\infty} Z_{n,k}^{\gamma}(\xi, \eta) = \xi^{n-k}\eta^k. \tag{4.4}$$

From (2.14) and (2.15) we get the special cases

$$Z_{n,k}^{-1/2}(x+y, xy) = (1+\delta_{n,k})^{-1}(x^n y^k + x^k y^n), \tag{4.5}$$

$$Z_{n,k}^{1/2}(x+y, xy) = (x-y)^{-1}(x^{n+1}y^k - x^k y^{n+1}). \tag{4.6}$$

From (2.17) we derive

$$Z_{n,k}^{\gamma}(x+y, xy) = \frac{(n-k)!}{(\gamma+\frac{1}{2})_{n-k}} \sum_{i=0}^{n-k} \frac{(\gamma+\frac{1}{2})_i(\gamma+\frac{1}{2})_{n-k-i}}{i!(n-k-i)!} x^{n-i} y^{k+i}. \tag{4.7}$$

Note also the boundary values

$$Z_{n,k}^{\gamma}(\xi, 0) = \begin{cases} \xi^n & \text{if } k = 0, \\ 0 & \text{if } k > 0, \end{cases} \tag{4.8}$$

$$Z_{n,k}^{\gamma}(\xi, \tfrac{1}{4}\xi^2) = \frac{(2\gamma+1)_{n-k}}{(\gamma+\frac{1}{2})_{n-k}} (\tfrac{1}{2}\xi)^{n+k}. \tag{4.9}$$

In view of (4.3) we may conclude that any polynomial

$$P(\xi, \eta) := \sum_{l=0}^{n} \sum_{m=l}^{n} c_{m,l}\xi^{m-l}\eta^l$$

has a unique expansion

$$P(\xi, \eta) = \sum_{l=0}^{n} \sum_{m=l}^{n} c_{m,l}^{\gamma} Z_{m,l}^{\gamma}(\xi, \eta)$$

for each $\gamma > -1$. This can be considered as a generalized power series expansion.

Note that $c_{m,0}^{\gamma} = c_{m,0}$ by (4.8).

In particular, let us define the coefficients $c_{n,k;m,l}^{\alpha,\beta,\gamma}$ by

DEFINITION 4.2.

$$R_{n,k}^{\alpha,\beta,\gamma}(\xi, \eta) = \sum_{l=0}^{n} \sum_{m=l}^{n} c_{n,k;m,l}^{\alpha,\beta,\gamma} Z_{m,l}^{\gamma}(\xi, \eta). \tag{4.10}$$

We claim that the generalized power series expansion (4.10) is a suitable analogue of the ordinary power series expansion (2.4) for Jacobi polynomials. This also justifies the introduction of the new coordinates ξ, η in § 3. Below we give a number of arguments for considering the expansion (4.10).

(a) It follows from (3.4), (3.5) and (2.4) that $R_{n,k}^{\alpha,\beta,-1/2}(x+y, xy)$ has a natural expansion in terms of $x^m y^l + x^l y^m$, $m \geqq l$, and, similarly, $R_{n,k}^{\alpha,\beta,1/2}(x+y, xy)$ in terms of $(x-y)^{-1}(x^{m+1}y^l - x^l y^{m+1})$, $m \geqq l$. In both cases the expansion coefficients can be given explicitly. By (4.5) and (4.6) this leads to the expansion (4.10) in the case $\gamma = \pm\frac{1}{2}$ and we obtain

$$\begin{aligned} c_{n,k;m,l}^{\alpha,\beta,-1/2} = \{&(-n)_m(-k)_l(n+\alpha+\beta+1)_m(k+\alpha+\beta+1)_l \\ &+(-k)_m(-n)_l(k+\alpha+\beta+1)_m(n+\alpha+\beta+1)_l\}\{2(\alpha+1)_m(\alpha+1)_l m!\,l!\}^{-1}, \end{aligned} \tag{4.11}$$

$$\begin{aligned} c_{n,k;m,l}^{\alpha,\beta,1/2} = -\{&(-n-1)_{m+1}(-k)_l(n+\alpha+\beta+2)_{m+1}(k+\alpha+\beta+1)_l \\ &-(-k)_{m+1}(-n-1)_l(k+\alpha+\beta+1)_{m+1}(n+\alpha+\beta+2)_l\} \\ &\cdot\{(n-k+1)(n+k+\alpha+\beta+2)(\alpha+2)_m(\alpha+1)_l(m+1)!\,l!\}^{-1}. \end{aligned} \tag{4.12}$$

(b) It was pointed out in Koornwinder [25, § 4.4] that

$$R_{n,n}^{\alpha,\beta,0}(x+y, xy) = {}_2F_1\left(-n, n+\alpha+\beta+\tfrac{3}{2}; \alpha+\tfrac{3}{2}; \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix}\right), \tag{4.13}$$

where ${}_2F_1(a, b; c; X)$ is the hypergeometric function of matrix argument X which was introduced by Herz [17]. Constantine [10] proved that there are natural power series expansions of such hypergeometric functions in terms of so-called zonal polynomials which (in the (2×2) case) are spherical functions on $GL(2, \mathbb{R})/O(2)$ belonging to finite dimensional irreducible representations of $GL(2, \mathbb{R})$. These zonal polynomials were introduced by James [18]. Furthermore, James [19, (7.9)] showed that in the (2×2) case these zonal polynomials coincide up to a constant factor with our polynomials $Z_{m,l}^0(x+y, xy)$. By using formula (25) in Constantine [10] it follows that

$$\begin{aligned} &{}_2F_1\left(a, b; c; \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix}\right) \\ &\quad = \sum_{m=0}^{\infty}\sum_{l=0}^{\infty} \frac{(a)_m(a-\frac{1}{2})_l(b)_m(b-\frac{1}{2})_l(\frac{3}{2})_{m-l}}{(c)_m(c-\frac{1}{2})_l(\frac{3}{2})_m l!(m-l)!} Z_{m,l}^0(x+y, xy). \end{aligned} \tag{4.14}$$

Now (4.13) and (4.14) together give

$$c_{n,n;m,l}^{\alpha,\beta,0} = \frac{(-n)_m(-n-\frac{1}{2})_l(n+\alpha+\beta+\frac{3}{2})_m(n+\alpha+\beta+1)_l(\frac{3}{2})_{m-l}}{(\alpha+\frac{3}{2})_m(\alpha+1)_l(\frac{3}{2})_m l!(m-l)!}. \tag{4.15}$$

(c) It can be proved that there is an interpretation of the polynomials $R_{n,k}^{(q-3)/2,(d-q-3)/2,0}(\xi, \eta)$ as so-called intertwining functions on the group $O(d)$, which are right invariant with respect to $O(2)\times O(d-2)$, left invariant with respect to $O(q)\times O(d-q)$, and which belong to some irreducible representation of $O(d)$. In particular, for $q=2$ we obtain the spherical functions on the Grassmann manifold $O(d)/O(2)\times O(d-2)$. According to James and Constantine [20] group theoretic considerations give a motivation for expanding these intertwining functions in terms of zonal polynomials. In particular, it follows from James and Constantine [20, (15.4)] that

$$c_{n,0;m,l}^{\alpha,\beta,0} = \begin{cases} \dfrac{(-n)_m(n+\alpha+\beta+2)_m(\frac{1}{2})_m}{(\alpha+\frac{3}{2})_m m!\,m!} & \text{if } l=0, \\ 0 & \text{if } l\neq 0, \end{cases} \tag{4.16}$$

for integer or half-integer α and β.

(d) It will be proved in § 7 that

$$\lim_{\beta\to\infty} \frac{R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)}{R_{n,k}^{\alpha,\beta,\gamma}(2,1)} = \frac{Z_{n,k}^{\gamma}(\xi,\eta)}{Z_{n,k}^{\gamma}(2,1)}. \tag{4.17}$$

Note that the pair of formulas (4.10) and (4.17) is analogous with the formulas (2.4) and (2.21) for Jacobi polynomials.

A final motivation for considering the expansion (4.10) is given by the following differentiation formula, which is easily verified.

$$\begin{aligned} &D_-^{\gamma} Z_{n,0}^{\gamma}(\xi,\eta) = 0, \\ &D_-^{\gamma} Z_{n,k}^{\gamma}(\xi,\eta) = \tfrac{1}{4}k(n+\gamma+\tfrac{1}{2}) Z_{n-1,k-1}^{\gamma}(\xi,\eta) \quad \text{if } k>0. \end{aligned} \tag{4.18}$$

On comparing this result with (3.13) we obtain the recurrence relation

$$c_{n,k;m,l}^{\alpha,\beta,\gamma} = \frac{k(k+\alpha+\beta+1)(n+\gamma+\frac{1}{2})(n+\alpha+\beta+\gamma+\frac{3}{2})}{l(\alpha+1)(m+\gamma+\frac{1}{2})(\alpha+\gamma+\frac{3}{2})} c_{n-1,k-1;m-1,l-1}^{\alpha+1,\beta+1,\gamma}, \quad l>0, \quad k>0. \tag{4.19}$$

Since $D_-^{\gamma} R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta) = 0$, it also follows that

$$R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta) = \sum_{m=0}^{n} c_{n,0;m,0}^{\alpha,\beta,\gamma} Z_{m,0}^{\gamma}(\xi,\eta), \tag{4.20}$$

i.e. $c_{n,0;m,l}^{\alpha,\beta,\gamma} = 0$ if $l>0$. Formulas (4.19) and (4.20) together imply:

THEOREM 4.3. $c_{n,k;m,l}^{\alpha,\beta,\gamma} \neq 0$ *only if* $l \leqq k$ (cf. Fig. 3).

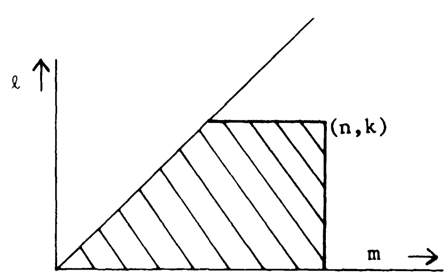

FIG. 3

Remark 4.4. Theorem 4.3 together with formula (4.2) provides a new proof of Theorem 3.2.

Remark 4.5. It follows from (4.10) and (4.8) that

$$R_{n,k}^{\alpha,\beta,\gamma}(\xi,0) = \sum_{m=0}^{n} c_{n,k;m,0}^{\alpha,\beta,\gamma} \xi^m. \tag{4.21}$$

Hence, in view of (4.19), we know the general coefficients $c_{n,k;m,l}^{\alpha,\beta,\gamma}$ as soon as we know the explicit power series expansion (4.21) of the boundary value $R_{n,k}^{\alpha,\beta,\gamma}(\xi,0)$ for all values of α,β,γ,n,k. In particular, we know the coefficients $c_{n,n;m,l}^{\alpha,\beta,\gamma}$ as soon as we know the explicit power series expansion of $R_{n,n}^{\alpha,\beta,\gamma}(\xi,0)$ for all values of α,β,γ,n. This power series expansion will be obtained in § 5.

Remark 4.6. It follows from (4.20) and (4.9) that

$$R_{n,0}^{\alpha,\beta,\gamma}(\xi,\tfrac{1}{4}\xi^2)=\sum_{m=0}^{n} c_{n,0;m,0}^{\alpha,\beta,\gamma}\frac{(2\gamma+1)_m}{(\gamma+\frac{1}{2})_m}(\tfrac{1}{2}\xi)^m. \tag{4.22}$$

Hence, the general expansion (4.20) of $R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta)$ is known as soon as we know the explicit power series expansion (4.22) of the boundary value $R_{n,0}^{\alpha,\beta,\gamma}(\xi,\frac{1}{4}\xi^2)$. This power series expansion will also be obtained in § 5.

5. Some boundary values. In this section it will be shown that the polynomials $R_{n,n}^{\alpha,\beta,\gamma}(\xi,\eta)$ become Jacobi polynomials on the boundary lines $\eta=0$ and $1-\xi+\eta=0$, and that $R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta)$ can be expressed as a Jacobi polynomial on the parabola $\xi^2-4\eta=0$. In the case of general degree (n,k) certain Jacobi expansions of the boundary values of $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ will be considered, for which the Fourier–Jacobi coefficients can be expressed in terms of the corresponding coefficients for $R_{n-k,0}^{\alpha+k,\beta+k,\gamma}(\xi,0)$ and $R_{k,k}^{\alpha,\beta,\gamma+n-k}(\xi,\frac{1}{4}\xi^2)$, respectively.

The key for deriving these results is the following lemma.

LEMMA 5.1.

(a) *On the line $\eta=0$ the second order partial differential operator $D_+^{\alpha,\beta,\gamma}$ reduces to a first order ordinary differential operator involving only derivatives $\partial_\xi=d/d\xi$, which is given by*

$$D_+^{\alpha,\beta,\gamma}|_{\eta=0}=4(\alpha+1)\xi^{-(\alpha+\gamma+1/2)}(1-\xi)^{-\beta}\frac{d}{d\xi}\circ\xi^{\alpha+\gamma+3/2}(1-\xi)^{\beta+1}. \tag{5.1}$$

(b) *On the parabola $\xi^2-4\eta=0$ the operator $E_+^{\alpha,\beta,\gamma}$ reduces to a first order differential operator involving only derivatives $\partial_\xi+\frac{1}{2}\xi\,\partial_\eta=d/d\xi$, which is given by*

$$E_+^{\alpha,\beta,\gamma}|_{\xi^2-4\eta=0}=4(\gamma+1)\xi^{-(\alpha+\gamma+1/2)}(2-\xi)^{-(\beta+\gamma+1/2)}\frac{d}{d\xi}\circ\xi^{\alpha+\gamma+3/2}(2-\xi)^{\beta+\gamma+3/2}. \tag{5.2}$$

Proof. The proof follows immediately by substitution of $\eta=0$ in (3.11) and $\xi^2-4\eta=0$ in (3.12), respectively. □

Formulas (5.1), (5.2) and (2.3) imply that

$$D_+^{\alpha,\beta,\gamma}|_{\eta=0}R_{n-1}^{(\alpha+\gamma+3/2,\beta+1)}(1-2\xi)=4(\alpha+1)(\alpha+\gamma+\tfrac{3}{2})R_n^{(\alpha+\gamma+1/2,\beta)}(1-2\xi), \tag{5.3}$$

$$E_+^{\alpha,\beta,\gamma}|_{\xi^2-4\eta=0}R_{n-1}^{(\alpha+\gamma+3/2,\beta+\gamma+3/2)}(1-\xi)=8(\gamma+1)(\alpha+\gamma+\tfrac{3}{2})R_n^{(\alpha+\gamma+1/2,\beta+\gamma+1/2)}(1-\xi). \tag{5.4}$$

Now we can prove the important

THEOREM 5.2.

$$R_{n,n}^{\alpha,\beta,\gamma}(\xi,0)=R_n^{(\alpha+\gamma+1/2,\beta)}(1-2\xi), \tag{5.5}$$

$$R_{n,0}^{\alpha,\beta,\gamma}(\xi,\tfrac{1}{4}\xi^2)=R_n^{(\alpha+\gamma+1/2,\beta+\gamma+1/2)}(1-\xi), \tag{5.6}$$

$$\frac{R_{n,n}^{\alpha,\beta,\gamma}(\xi,\xi-1)}{R_{n,n}^{\alpha,\beta,\gamma}(2,1)}=R_n^{(\beta+\gamma+1/2,\alpha)}(2\xi-3). \tag{5.7}$$

Proof. Comparison of (3.14) with (5.3) and of (3.16) with (5.4) and complete induction with respect to n results in (5.5) and (5.6). The boundary value of $R_{n,n}^{\alpha,\beta,\gamma}(\xi,\eta)$ for $1-\xi+\eta=0$ follows from (5.5) and (3.18). □

Next we will consider Jacobi expansions for the polynomials $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ on the boundary curves $\eta=0$ and $\xi^2-4\eta=0$. Let us define the coefficients $a_{n,k;m}^{\alpha,\beta,\gamma}$ and $b_{n,k;m}^{\alpha,\beta,\gamma}$ by

DEFINITION 5.3.

$$R_{n,k}^{\alpha,\beta,\gamma}(\xi,0)=\sum_m a_{n,k;m}^{\alpha,\beta,\gamma}R_m^{(\alpha+\gamma+1/2,\beta)}(1-2\xi), \tag{5.8}$$

$$R_{n,k}^{\alpha,\beta,\gamma}(\xi,\tfrac14\xi^2)=\sum_m b_{n,k;m}^{\alpha,\beta,\gamma}R_m^{(\alpha+\gamma+1/2,\beta+\gamma+1/2)}(1-\xi). \tag{5.9}$$

Formulas (5.3) and (3.14), (5.4) and (3.16) yield the following equalities for the coefficients $a_{n,k;m}^{\alpha,\beta,\gamma}(k>0)$ and $b_{n,k;m}^{\alpha,\beta,\gamma}(n>k)$, respectively:

$$a_{n,k;m}^{\alpha,\beta,\gamma}=\begin{cases}a_{n-1,k-1;m-1}^{\alpha+1,\beta+1,\gamma} & \text{if } m>0,\\ 0 & \text{if } m=0,\end{cases} \tag{5.10}$$

$$b_{n,k;m}^{\alpha,\beta,\gamma}=\begin{cases}b_{n-1,k;m-1}^{\alpha,\beta,\gamma+1} & \text{if } m>0,\\ 0 & \text{if } m=0.\end{cases} \tag{5.11}$$

Use of (5.10), (5.11) and complete induction with respect to k and $n-k$, respectively, results in

$$a_{n,k;m}^{\alpha,\beta,\gamma}\neq 0 \quad \text{only if } k\leqq m\leqq n, \tag{5.12}$$

$$b_{n,k;m}^{\alpha,\beta,\gamma}\neq 0 \quad \text{only if } n-k\leqq m\leqq n+k, \tag{5.13}$$

$$a_{n,k;m}^{\alpha,\beta,\gamma}=a_{n-k,0;m-k}^{\alpha+k,\beta+k,\gamma}, \tag{5.14}$$

$$b_{n,k;m}^{\alpha,\beta,\gamma}=b_{k,k;m-n+k}^{\alpha,\beta,\gamma+n-k}. \tag{5.15}$$

Thus we obtain

LEMMA 5.4. *We have*

$$R_{n,k}^{\alpha,\beta,\gamma}(\xi,0)=\sum_{m=k}^{n} a_{n-k,0;m-k}^{\alpha+k,\beta+k,\gamma}R_m^{(\alpha+\gamma+1/2,\beta)}(1-2\xi), \tag{5.16}$$

$$R_{n,k}^{\alpha,\beta,\gamma}(\xi,\tfrac14\xi^2)=\sum_{m=n-k}^{n+k} b_{k,k;m-n+k}^{\alpha,\beta,\gamma+n-k}R_m^{(\alpha+\gamma+1/2,\beta+\gamma+1/2)}(1-\xi), \tag{5.17}$$

$$\frac{R_{n,k}^{\alpha,\beta,\gamma}(\xi,\xi-1)}{R_{n,k}^{\alpha,\beta,\gamma}(2,1)}=\sum_{m=k}^{n} a_{n-k,0;m-k}^{\beta+k,\alpha+k,\gamma}R_m^{(\beta+\gamma+1/2,\alpha)}(2\xi-3). \tag{5.18}$$

The last part of the lemma follows from (5.16) and (3.18).

Next we give some formulas for $a_{n,k;m}^{\alpha,\beta,\gamma}$ and $b_{n,k;m}^{\alpha,\beta,\gamma}$ in the case of special values of the parameters.

From (3.4) and (2.40) it follows that $b_{n,k;m}^{\alpha,\beta,-1/2}$ can be expressed in terms of the linearization coefficients $A_{n,k,m}^{\alpha,\beta,\gamma}$ of the Jacobi polynomials:

$$b_{n,k;m}^{\alpha,\beta,-1/2}=A_{n,k,m}^{(\alpha,\beta)}\omega_m^{(\alpha,\beta)}. \tag{5.19}$$

If $\alpha=\beta$ then application of the quadratic transformation formulas (3.19) and (2.9), (3.20) and (2.10), respectively, results in

$$b_{n+k,n-k;2m}^{\alpha,\alpha,-1/2}=a_{n,k;m}^{-1/2,-1/2,\alpha}, \tag{5.20}$$

$$b_{n+k+1,n-k;2m+1}^{\alpha,\alpha,-1/2}=a_{n,k;m}^{-1/2,1/2,\alpha}. \tag{5.21}$$

From (3.18) and (5.9) with $\alpha=\beta$ it follows that

$$b_{n,n;2m+1}^{\alpha,\alpha,\gamma}=b_{n+1,n;2m}^{\alpha,\alpha,\gamma}=0. \tag{5.22}$$

Combination of (5.19), (5.20), (5.21) and (5.22) gives an expression for the linearization coefficients of order (α,α) in terms of $a_{n,k;m}^{-1/2,\pm1/2,\alpha}$. In § 6 we will derive the explicit values of the coefficients $a_{n,k;m}^{\alpha,\beta,\gamma}$ and thus we will find a new derivation of the linearization coefficients for the Gegenbauer polynomials.

It follows from the quadratic transformation formulas (3.19), (3.20), (2.9) and (2.10) that

$$b_{n,n;2m}^{\alpha,\alpha,\gamma}=a_{n,0;m}^{\gamma,-1/2,\alpha}, \tag{5.23}$$

$$b_{n,n;m}^{\alpha,-1/2,\gamma}=a_{2n,0;m}^{\gamma,\gamma,\alpha}, \tag{5.24}$$

$$b_{n,n;2m}^{\alpha,\alpha,\gamma+1}=a_{n,0;m}^{\gamma,1/2,\alpha}. \tag{5.25}$$

For the proof of (5.25) we used (5.11) once.

From Lemma 5.4 and Lemma 2.5 we can derive a corollary about the number of zeros of $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ on the boundary.

COROLLARY 5.5.

(a) $R_{n,k}^{\alpha,\beta,\gamma}(\xi,0)$ *has at least* k *zeros of odd multiplicity for* $\xi\in(0,1)$.

(b) $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\xi-1)$ *has at least* k *zeros of odd multiplicity for* $\xi\in(1,2)$.

(c) $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\frac{1}{4}\xi^2)$ *has at least* $n-k$ *zeros of odd multiplicity for* $\xi\in(0,2)$.

6. Expansion of the polynomials $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ in terms of James-type zonal polynomials. In this section we will derive the explicit value of the coefficients $c_{n,k;m,l}^{\alpha,\beta,\gamma}$ in formula (4.10) giving the expansion of $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ in terms of the James-type zonal polynomials $Z_{m,l}^{\gamma}(\xi,\eta)$. We will proceed in the following way. From the boundary value $R_{n,0}^{\alpha,\beta,\gamma}(\xi,\frac{1}{4}\xi^2)$ we obtain the coefficients $c_{n,0;m,0}^{\alpha,\beta,\gamma}$, and hence the boundary value $R_{n,0}^{\alpha,\beta,\gamma}(\xi,0)$. By rewriting $R_{n,0}^{\alpha,\beta,\gamma}(\xi,0)$ as a Jacobi series we derive the coefficients $a_{n,0;m}^{\alpha,\beta,\gamma}$ defined by (5.8). This also gives the coefficients $a_{n,k;m}^{\alpha,\beta,\gamma}$. Next the Jacobi series of $R_{n,k}^{\alpha,\beta,\gamma}(\xi,0)$ can be rewritten as a power series and we obtain the coefficients $c_{n,k;m,0}^{\alpha,\beta,\gamma}$. Finally $c_{n,k;m,l}^{\alpha,\beta,\gamma}$ can be expressed in terms of $c_{n-l,k-l;m-l,0}^{\alpha+l,\beta+l,\gamma}$.

At the end of this section several interesting corollaries will be discussed. We mention the expression of $R_{n,0}^{\alpha,\beta,\gamma}(x+y,xy)$ as a generalized hypergeometric function in the two variables x and y, the expression of $R_{n,k}^{\alpha,\beta,\gamma}(1,0)$ in terms of a ${}_3F_2$-function of argument 1, and a new derivation of the linearization coefficients for Gegenbauer polynomials.

Let us consider $R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta)$. Combination of (5.6), (4.22) and (2.4) results in

$$c_{n,0;m,0}^{\alpha,\beta,\gamma}=\frac{(-n)_m(n+\alpha+\beta+2\gamma+2)_m(\gamma+\frac{1}{2})_m}{(\alpha+\gamma+\frac{3}{2})_m(2\gamma+1)_m m!}. \tag{6.1}$$

So we have the first explicit expansion (cf. (4.20)):

$$R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta)=\sum_{m=0}^{n}\frac{(-n)_m(n+\alpha+\beta+2\gamma+2)_m(\gamma+\frac{1}{2})_m}{(\alpha+\gamma+\frac{3}{2})_m(2\gamma+1)_m m!}Z_{m,0}^{\gamma}(\xi,\eta). \tag{6.2}$$

Remark 6.1. If $\gamma=-\frac{1}{2}$ then the right hand side of formula (6.2) has to be interpreted as the limit case for $\gamma\to-\frac{1}{2}$. A similar interpretation has to be used on many other places.

Similar to (6.1), it follows from (5.5), (4.21) and (2.4) that

$$c_{n,n;m,0}^{\alpha,\beta,\gamma}=\frac{(-n)_m(n+\alpha+\beta+\gamma+\frac{3}{2})_m}{(\alpha+\gamma+\frac{3}{2})_m m!}. \tag{6.3}$$

Hence, by (4.19) we have

$$c_{n,n;m,l}^{\alpha,\beta,\gamma}=\frac{(-n)_m(-n-\gamma-\frac{1}{2})_l(n+\alpha+\beta+\gamma+\frac{3}{2})_m(n+\alpha+\beta+1)_l(\gamma+\frac{3}{2})_{m-l}}{(\alpha+\gamma+\frac{3}{2})_m(\alpha+1)_l(\gamma+\frac{3}{2})_m l!(m-l)!} \tag{6.4}$$

and thus the expansion (4.10) for the polynomials $R_{n,n}^{\alpha,\beta,\gamma}(\xi,\eta)$.

LEMMA 6.2. *The power series expansion of* $R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta)$ *for* $\eta=0$ *is given by*

$$R_{n,0}^{\alpha,\beta,\gamma}(\xi,0)={}_3F_2\left(\begin{matrix}-n,\, n+\alpha+\beta+2\gamma+2,\, \gamma+\frac{1}{2};\\ \alpha+\gamma+\frac{3}{2},\, 2\gamma+1;\end{matrix}\ \xi\right). \tag{6.5}$$

Proof. The proof follows immediately from (4.21) and (6.2). □

LEMMA 6.3. *The coefficients* $a_{n,0;m}^{\alpha,\beta,\gamma}$ *in the Jacobi expansion* (5.8) *of* $R_{n,0}^{\alpha,\beta,\gamma}(\xi,0)$ *are given by*

$$a_{n,0;m}^{\alpha,\beta,\gamma}=\frac{n!(n+\alpha+\beta+2\gamma+2)_m(m+\alpha+\beta+2)_{n-m}(\gamma+\frac{1}{2})_m(\gamma+\frac{1}{2})_{n-m}}{(2\gamma+1)_n(m+\alpha+\beta+\gamma+\frac{3}{2})_m(2m+\alpha+\beta+\gamma+\frac{5}{2})_{n-m}m!(n-m)!}. \tag{6.6}$$

Proof. It follows from (6.3), (2.6) and (5.8) that

$$a_{n,0;m}^{\alpha,\beta,\gamma}=\frac{(-1)^m(-n)_m(n+\alpha+\beta+2\gamma+2)_m(\gamma+\frac{1}{2})_m}{(m+\alpha+\beta+\gamma+\frac{3}{2})_m(2\gamma+1)_m m!}$$
$$\cdot\,{}_3F_2\left(\begin{matrix}-n+m,\, n+m+\alpha+\beta+2\gamma+2,\, m+\gamma+\frac{1}{2};\\ 2m+\alpha+\beta+\gamma+\frac{5}{2},\, m+2\gamma+1;\end{matrix}\ 1\right),$$

which can be evaluated by using (2.42). □

THEOREM 6.4. *The explicit form of the Jacobi expansion* (5.8) *is*

$$\begin{aligned}R_{n,k}^{\alpha,\beta,\gamma}(\xi,0)&=\frac{(n-k)!}{(2\gamma+1)_{n-k}}\sum_{m=k}^{n}\frac{(n+k+\alpha+\beta+2\gamma+2)_{m-k}}{(m+k+\alpha+\beta+\gamma+\frac{3}{2})_{m-k}}\\ &\quad\cdot\frac{(m+k+\alpha+\beta+2)_{n-m}(\gamma+\frac{1}{2})_{m-k}(\gamma+\frac{1}{2})_{n-m}}{(2m+\alpha+\beta+\gamma+\frac{5}{2})_{n-m}(m-k)!(n-m)!}R_m^{(\alpha+\gamma+1/2,\beta)}(1-2\xi).\end{aligned} \tag{6.7}$$

Note that

$$a_{n,k;m}^{\alpha,\beta,\gamma}>0\quad\text{if } \gamma>-\tfrac{1}{2}. \tag{6.8}$$

Proof. Use formulas (5.16) and (6.6). □

THEOREM 6.5. *The coefficients* $c_{n,k;m,0}^{\alpha,\beta,\gamma}$ *in the power series expansion* (4.10) *of* $R_{n,k}^{\alpha,\beta,\gamma}(\xi,0)$ *are given by*

(6.9)
$$c_{n,k;m,0}^{\alpha,\beta,\gamma}=\frac{(n+\alpha+\beta+\gamma+\frac{3}{2})_m(-n)_m}{(\alpha+\gamma+\frac{3}{2})_m m!}\,{}_4F_3\left(\begin{matrix}-m,\, -n+k,\, -n-k-\alpha-\beta-1,\, \gamma+\frac{1}{2};\\ -n,\, -n-m-\alpha-\beta-\gamma-\frac{1}{2},\, 2\gamma+1;\end{matrix}\ 1\right).$$

Proof. We will give two different proofs.

(a) It follows from (6.7) and (2.4) that

$$c_{n,k;m,0}^{\alpha,\beta,\gamma}=\frac{(\gamma+\frac{1}{2})_{n-k}(n+k+\alpha+\beta+2\gamma+2)_{n-k}(n+\alpha+\beta+\gamma+\frac{3}{2})_m(-n)_m}{(2\gamma+1)_{n-k}(n+k+\alpha+\beta+\gamma+\frac{3}{2})_{n-k}(\alpha+\gamma+\frac{3}{2})_m m!}$$
$$\begin{aligned}\cdot\,{}_7F_6(&-n+k,\, -n+m,\, \gamma+\tfrac{1}{2},\, -n-k-\alpha-\beta-1,\, -n-\alpha-\beta-\gamma-\tfrac{1}{2},\\ &-2n-\alpha-\beta-\gamma-\tfrac{3}{2},\, -n-\tfrac{1}{2}\alpha-\tfrac{1}{2}\beta-\tfrac{1}{2}\gamma+\tfrac{1}{4};\\ &-n,\, -n+k-\gamma+\tfrac{1}{2},\, -n-k-\alpha-\beta-\gamma-\tfrac{1}{2},\, -n-m-\alpha-\beta-\gamma-\tfrac{1}{2},\\ &-2n-\alpha-\beta-2\gamma-1,\, -n-\tfrac{1}{2}\alpha-\tfrac{1}{2}\beta-\tfrac{1}{2}\gamma-\tfrac{3}{4};\, 1).\end{aligned}$$

By using a result of Whipple (cf. Slater [27, (2.4.1.1)] this well-poised terminating ${}_7F_6$ can be rewritten as a Saalschützian terminating ${}_4F_3$ and the theorem follows.

(b) Combination of (3.14), (5.1) and (4.21) gives the recurrence relation

$$c_{n,k;m,0}^{\alpha,\beta,\gamma}=\frac{m+\alpha+\gamma+\frac{3}{2}}{\alpha+\gamma+\frac{3}{2}}c_{n-1,k-1;m,0}^{\alpha+1,\beta+1,\gamma}-\frac{m+\alpha+\beta+\gamma+\frac{3}{2}}{\alpha+\gamma+\frac{3}{2}}c_{n-1,k-1;m-1,0}^{\alpha+1,\beta+1,\gamma} \tag{6.10}$$

for $n \geqq k > 0$, $n \geqq m > 0$. For $k = 0$ the ${}_4F_3$ in (6.9) becomes a terminating Saalschützian ${}_3F_2$ which can be evaluated by using (2.42). In view of (6.5) the theorem turns out to be true for $k = 0$. Clearly, the theorem is true for $m = 0$. Using (6.10) we can now prove the general case of (6.9) by complete induction with respect to k. □

COROLLARY 6.6. *We have the expansion*

$$R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta) = \sum_{l=0}^{k} \sum_{m=l}^{n} c_{n,k;m,l}^{\alpha,\beta,\gamma} Z_{m,l}^{\gamma}(\xi,\eta), \tag{6.11}$$

where

$$\begin{aligned} c_{n,k;m,l}^{\alpha,\beta,\gamma} = {} & \frac{(-k)_l(-n)_m(-n-\gamma-\frac{1}{2})_l(n+\alpha+\beta+\gamma+\frac{3}{2})_m}{(-n)_l(\alpha+\gamma+\frac{3}{2})_m(\alpha+1)_l(\gamma+\frac{3}{2})_m} \\ & \cdot \frac{(k+\alpha+\beta+1)_l(\gamma+\frac{3}{2})_{m-l}}{l!(m-l)!} \\ & \cdot {}_4F_3\left(\begin{matrix} -m+l, -n+k, -n-k-\alpha-\beta-1, \gamma+\frac{1}{2}; \\ -n+l, -n-m-\alpha-\beta-\gamma-\frac{1}{2}, 2\gamma+1; \end{matrix}\ 1\right). \end{aligned} \tag{6.12}$$

In this expansion $Z_{m,l}^{\gamma}(\xi,\eta)$ is defined by (4.1).

Proof. By using complete induction with respect to k the result follows from (6.9) and (4.19). □

Remark 6.7. In a number of special cases of $m, l, n, k, \alpha, \beta, \gamma$ the expression (6.12) can be simplified. If one of the equalities $m = n$, $l = k$, $m = l$, $k = 0$ or $n = k$ holds, then the coefficient $c_{n,k;m,l}^{\alpha,\beta,\gamma}$ can be written as a quotient of products of gamma functions depending linearly on $m, l, n, k, \alpha, \beta, \gamma$. If $\gamma = \pm\frac{1}{2}$ then we get back (4.11) and (4.12).

We would like to write $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ as a linear combination of elementary expressions $(a/b)\xi^{m-l}\eta^l$, where a and b are products of gamma functions depending linearly on $m, l, n, k, \alpha, \beta, \gamma$. The best possible result would be a double sum, which indeed can be obtained for $R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta)$ and for $R_{n,n}^{\alpha,\beta,\gamma}(\xi,\eta)$ (see § 7). However, (6.11) expresses $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ as a quadruple sum of elementary terms. It is not clear to the authors how this can be simplified.

We conclude this section with a number of corollaries to the results earlier obtained in this section.

Combination of (6.2) and (4.7) gives:

COROLLARY 6.8. *We have*

$$R_{n,0}^{\alpha,\beta,\gamma}(x+y,xy) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(-n)_{i+j}(n+\alpha+\beta+2\gamma+2)_{i+j}(\gamma+\frac{1}{2})_i(\gamma+\frac{1}{2})_j}{(\alpha+\gamma+\frac{3}{2})_{i+j}(2\gamma+1)_{i+j}i!j!} x^i y^j. \tag{6.13}$$

With the use of the notation of Burchnall and Chaundy [7, § 1] this becomes

$$R_{n,0}^{\alpha,\beta,\gamma}(x+y,xy) = F\left(\begin{matrix} -n, n+\alpha+\beta+2\gamma+2 : \gamma+\frac{1}{2}; \gamma+\frac{1}{2}; \\ \alpha+\gamma+\frac{3}{2}, 2\gamma+1 \quad : \quad ; \quad ; \end{matrix}\ x, y\right),$$

a hypergeometric series in two variables of order three (cf. Erdélyi [12, § 5.7]). According to Carlson [9, (1.8)] it follows from (6.13) that

$$R_{n,0}^{\alpha,\beta,\gamma}(x+y,xy) = \mathcal{R}_n\left(n+\alpha+\beta+2\gamma+2, -n-\beta-\gamma-\tfrac{1}{2}; \begin{bmatrix} 1-x & 1-y \\ 1 & 1 \end{bmatrix}; \gamma+\tfrac{1}{2}, \gamma+\tfrac{1}{2}\right),$$

where the function $\mathcal{R}_t(\mu,\mu'; Y; \nu,\nu')$ is defined by Carlson [8, § 2].

COROLLARY 6.9. *The value of* $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ *in the vertex* $(1,0)$ *is given by*

$$R_{n,k}^{\alpha,\beta,\gamma}(1,0)=\frac{(-1)^k(\beta+1)_k}{(\alpha+\gamma+\frac{3}{2})_k}\,{}_3F_2\left(\begin{matrix}-n+k,\, n+k+\alpha+\beta+2\gamma+2,\, \gamma+\frac{1}{2};\\ k+\alpha+\gamma+\frac{3}{2},\, 2\gamma+1;\end{matrix}\ 1\right). \tag{6.14}$$

Proof. With (3.14) restricted to $(\xi,\eta)=(1,0)$ and from the use of (3.11) with $(\xi,\eta)=(1,0)$ it follows that

$$R_{n,k}^{\alpha,\beta,\gamma}(1,0)=-\frac{\beta+1}{\alpha+\gamma+\frac{3}{2}}R_{n-1,k-1}^{\alpha+1,\beta+1,\gamma}(1,0). \tag{6.15}$$

The corollary follows by iteration of this result and by using (6.5). □

Two special cases of (6.14) are $n=k$ and $\alpha=\beta$. In these cases we have, respectively,

$$R_{n,n}^{\alpha,\beta,\gamma}(1,0)=\frac{(-1)^n(\beta+1)_n}{(\alpha+\gamma+\frac{3}{2})_n}, \tag{6.16}$$

$$\begin{aligned} R_{n+k,n-k}^{\alpha,\alpha,\gamma}(1,0)&=(-1)^{n-k}\frac{(\alpha+1)_n(\frac{1}{2})_k}{(\alpha+\gamma+\frac{3}{2})_n(\gamma+1)_k},\\ R_{n+k+1,n-k}^{\alpha,\alpha,\gamma}(1,0)&=0. \end{aligned} \tag{6.17}$$

Formula (6.17) can be proved by application of Watson's formula (cf. Slater [27, (2.3.3.13)] or directly from (3.19), (3.20) and (3.17).

COROLLARY 6.10. *If* $\alpha\geqq\beta$, $\gamma\geqq-\frac{1}{2}$ *and* $\max(\alpha,\beta+\gamma+\frac{1}{2})\geqq-\frac{1}{2}$ *then*

$$\begin{aligned} |R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)|\leqq 1 \quad &\textit{for } \eta=0,\quad 0\leqq\xi\leqq 1,\\ &\textit{or } 1-\xi+\eta=0,\quad 1\leqq\xi\leqq 2. \end{aligned} \tag{6.18}$$

Proof. We use (6.7) together with the nonnegativity of the coefficients for $\gamma\geqq-\frac{1}{2}$ and the inequalities for Jacobi polynomials (cf. Theorem 2.2). The inequalities for α,β,γ imply that $\alpha+\gamma+\frac{1}{2}\geqq\beta$ and $\alpha+\gamma+\frac{1}{2}\geqq-\frac{1}{2}$. Hence

$$|R_{n,k}^{\alpha,\beta,\gamma}(\xi,0)|\leqq 1\quad \text{for } 0\leqq\xi\leqq 1.$$

In particular, $|R_{n,k}^{\alpha,\beta,\gamma}(1,0)|\leqq 1$. Since $\alpha\geqq\beta$ we also have $|R_{n,k}^{\alpha,\beta,\gamma}(2,1)|\leqq 1$ by (3.18).

Again by (3.18) we have

$$\begin{aligned} \left|\frac{R_{n,k}^{\alpha,\beta,\gamma}(2-\xi,1-\xi+\eta)}{R_{n,k}^{\alpha,\beta,\gamma}(2,1)}\right| &= |R_{n,k}^{\beta,\alpha,\gamma}(\xi,0)|\\ &\leqq \max\{|R_{n,k}^{\beta,\alpha,\gamma}(0,0)|,|R_{n,k}^{\beta,\alpha,\gamma}(1,0)|\}, \end{aligned}$$

since $\max(\alpha,\beta+\gamma+\frac{1}{2})\geqq-\frac{1}{2}$. Hence

$$|R_{n,k}^{\alpha,\beta,\gamma}(2-\xi,1-\xi+\eta)|\leqq\max\{|R_{n,k}^{\alpha,\beta,\gamma}(2,1)|,|R_{n,k}^{\alpha,\beta,\gamma}(1,0)|\}\leqq 1. \quad □$$

COROLLARY 6.11. *If* $\alpha\geqq-\frac{1}{2}$, $\alpha\geqq\beta$, $\gamma\geqq-\frac{1}{2}$ *and if one of the equalities* $\alpha=\beta$, $\beta=-\frac{1}{2}$ *or* $\gamma=-\frac{1}{2}$ *holds then*

$$|R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)|\leqq 1,\quad (\xi,\eta)\in\partial\Omega. \tag{6.19}$$

Proof. Use Corollary 6.10, formula (5.9) or (5.17), the nonnegativity of $b_{n,k;m}^{\alpha,\beta,\gamma}$ in the cases given in the corollary (cf. (5.22), (5.23), (5.24), (5.19) and Theorem 2.4), and the inequalities for the Jacobi polynomials (Theorem 2.2). □

Remark 6.12. We need the restricting equalities $\alpha=\beta$, $\beta=-\frac{1}{2}$ or $\gamma=-\frac{1}{2}$ in Corollary 6.11 because in other cases the nonnegativity of $b_{n,k;m}^{\alpha,\beta,\gamma}$ is not yet proved. It is the authors' hypothesis that the coefficients $b_{n,k;m}^{\alpha,\beta,\gamma}$ are positive for $\alpha\geqq-\frac{1}{2}$, $\alpha\geqq\beta$ and $\gamma\geqq-\frac{1}{2}$. If this is true then formula (6.19) would hold for all α,β,γ such that $\alpha\geqq-\frac{1}{2}$, $\alpha\geqq\beta$, $\gamma\geqq-\frac{1}{2}$.

Remark 6.13. Combination of formulas (6.7), (5.19), (5.20), (5.21) and (5.22) results in a new proof for the linearization coefficients $A_{n,k,m}^{(\alpha,\alpha)}$ for the Gegenbauer polynomials (cf. (2.40) and (2.41)).

We can use (6.11) with coefficients given by (6.12) in order to derive the following pair of differential recurrence relations, which are the analogues of (2.24), (2.25) for Jacobi polynomials.

COROLLARY 6.14. *We have*

$$\eta^{1-\alpha}D_-^\gamma\circ\eta^\alpha R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)=\tfrac{1}{4}\alpha(\alpha+\gamma+\tfrac{1}{2})R_{n,k}^{\alpha-1,\beta+1,\gamma}(\xi,\eta), \tag{6.20}$$

$$(1-\xi+\eta)^{-\beta}D_-^\gamma\circ(1-\xi+\eta)^{\beta+1}\frac{R_{n,k}^{\alpha-1,\beta+1,\gamma}(\xi,\eta)}{R_{n,k}^{\alpha-1,\beta+1,\gamma}(2,1)} =\tfrac{1}{4}(\beta+1)(\beta+\gamma+\tfrac{3}{2})\frac{R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)}{R_{n,k}^{\alpha,\beta,\gamma}(2,1)}. \tag{6.21}$$

Proof. Formula (4.18) can be extended to

$$\eta^{1-\alpha}D_-^\gamma\circ\eta^\alpha Z_{n,k}^\gamma(\xi,\eta)=\tfrac{1}{4}(k+\alpha)(n+\alpha+\gamma+\tfrac{1}{2})Z_{n,k}^\gamma(\xi,\eta). \tag{6.22}$$

Substitution of (6.22) in (6.11) gives (6.20). Combination of (6.20) and (3.18) gives (6.21). □

7. Another expansion of the polynomials $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ and their relation with Appell's function F_4. In this section we will consider an expansion of $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ in terms of the polynomials.

$$(1-\xi)^m R_l^{(\alpha,\beta)}(1+2(1-\xi)^{-1}\eta),\qquad m\geqq l. \tag{7.1}$$

These polynomials play a similar role with respect to the operator $E_-^{\alpha,\beta}$ as the James-type zonal polynomials do with respect to D_-^γ. In particular, it will be proved that the expansion of $R_{n,n}^{\alpha,\beta,\gamma}(\xi,\eta)$ only contains polynomials for which $m=l$. It will follow from this expansion that $R_{n,n}^{\alpha,\beta,\gamma}(\xi,\eta)$ can be expressed as an Appell function F_4, which seems to be a quite important result. This section will be concluded with an interpretation of $Z_{m,l}^\gamma(\xi,\eta)$ and the polynomials (7.1) as limit cases of $R_{m,l}^{\alpha,\beta,\gamma}(\xi,\eta)$ for $\beta\to\infty$, $\gamma\to\infty$, respectively.

Let us consider the polynomial (7.1). By (2.4) its power series expansion in $1-\xi$ and η equals

$$(1-\xi)^n R_k^{(\alpha,\beta)}(1+2(1-\xi)^{-1}\eta)=\sum_{i=0}^k\frac{(-k)_i(k+\alpha+\beta+1)_i}{(\alpha+1)_i i!}(1-\xi)^{n-i}(-\eta)^i. \tag{7.2}$$

On the boundary line $\eta=0$ the polynomial reduces to

$$(1-\xi)^n R_k^{(\alpha,\beta)}(1+2(1-\xi)^{-1}\eta)\Big|_{\eta=0}=(1-\xi)^n. \tag{7.3}$$

By the use of (7.2) the polynomial restricted to the axis of reflection $\xi=1$ becomes

$$(1-\xi)^n R_k^{(\alpha,\beta)}(1+2(1-\xi)^{-1}\eta)\Big|_{\xi=1}=\begin{cases}\dfrac{(n+\alpha+\beta+1)_n}{(\alpha+1)_n}\eta^n & \text{if } n=k,\\ 0 & \text{if } n>k.\end{cases} \tag{7.4}$$

In view of (7.2) we conclude that any polynomial in ξ and η has a unique expansion in terms of the polynomials (7.1) (α,β fixed). In particular, we will consider the expansion of $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ in terms of these polynomials.

DEFINITION 7.1. The coefficients $d_{n,k;m,l}^{\alpha,\beta,\gamma}$ are given by

$$R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)=\sum_{l=0}^{n}\sum_{m=l}^{n} d_{n,k;m,l}^{\alpha,\beta,\gamma}(1-\xi)^m R_l^{(\alpha,\beta)}(1+2(1-\xi)^{-1}\eta). \tag{7.5}$$

Remark 7.2. If $\xi=1$ then substitution of (7.4) in (7.5) results in

$$R_{n,k}^{\alpha,\beta,\gamma}(1-\eta)=\sum_{m=0}^{n} d_{n,k;m,m}^{\alpha,\beta,\gamma}\frac{(m+\alpha+\beta+1)_m}{(\alpha+1)_m}\eta^m. \tag{7.6}$$

The following theorem gives a motivation for considering the expansion (7.5).

THEOREM 7.3. *We have*

$$E_-^{\alpha,\beta}(1-\xi)^n R_k^{(\alpha,\beta)}(1+2(1-\xi)^{-1}\eta)$$
$$=\begin{cases}\tfrac12(n-k)(n+k+\alpha+\beta+1)(1-\xi)^{n-1}R_k^{(\alpha,\beta)}(1+2(1-\xi)^{-1}\eta) & \text{if } n>k,\\ 0 & \text{if } n=k.\end{cases} \tag{7.7}$$

Proof. Use (3.9) and (2.1). □

On comparing (7.7) with (7.5) and (3.15) we obtain the recurrence relation

$$d_{n,k;m,l}^{\alpha,\beta,\gamma}=\frac{(n-k)(n-k+2\gamma+1)(n+k+\alpha+\beta+1)(n+k+\alpha+\beta+2\gamma+2)}{4(\gamma+1)(\alpha+\gamma+\frac32)(m-l)(m+l+\alpha+\beta+1)}d_{n-1,k;m-1,l}^{\alpha,\beta,\gamma+1} \quad \text{if } m>l \text{ and } n>k. \tag{7.8}$$

Since $E_-^{\alpha,\beta}R_{n,n}^{\alpha,\beta,\gamma}(\xi,\eta)=0$, it also follows that

$$R_{n,n}^{\alpha,\beta,\gamma}(\xi,\eta)=\sum_{m=0}^{n} d_{n,n;m,m}^{\alpha,\beta,\gamma}(1-\xi)^m R_m^{(\alpha,\beta)}(1+2(1-\xi)^{-1}\eta). \tag{7.9}$$

THEOREM 7.4. *The coefficients $d_{n,k;m,l}^{\alpha,\beta,\gamma}$ in* (7.5) *are nonzero only if $m-l\leqq n-k$ and $m+l\leqq n+k$* (cf. Fig. 4).

Proof. The inequality $m+l\leqq n+k$ follows from (7.1) and Theorem 3.2. The inequality $m-l\leqq n-k$ is a consequence of (7.8) and (7.9). □

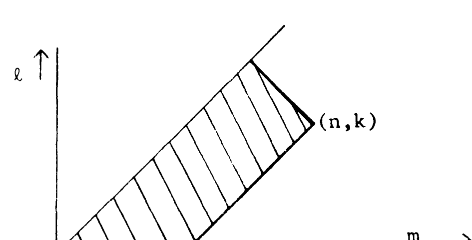

FIG. 4

In view of (7.6), (7.8) and Theorem 7.4, we obtain the coefficients $d^{\alpha,\beta,\gamma}_{n,k;m,l}$ as soon as we know the expansion of $R^{\alpha,\beta,\gamma+m-l}_{n-m+l,k}(1,\eta)$ as a power series in η. Here we restrict ourselves to the case $n=k$. It follows from (7.9) that

$$R^{\alpha,\beta,\gamma}_{n,n}(\xi,0)=\sum_{m=0}^{n} d^{\alpha,\beta,\gamma}_{n,n;m,m}(1-\xi)^m. \tag{7.10}$$

From (5.5), (2.7), (2.8) and (2.4) we know

$$\frac{R^{\alpha,\beta,\gamma}_{n,n}(\xi,0)}{R^{\alpha,\beta,\gamma}_{n,n}(1,0)}=\sum_{m=0}^{n}\frac{(-n)_m(n+\alpha+\beta+\gamma+\frac{3}{2})_m}{(\beta+1)_m m!}(1-\xi)^m. \tag{7.11}$$

Comparison of (7.10) and (7.11) yields (by the use of (6.16))

$$d^{\alpha,\beta,\gamma}_{n,n;m,m}=\frac{(-1)^n(\beta+1)_n(-n)_m(n+\alpha+\beta+\gamma+\frac{3}{2})_m}{(\alpha+\gamma+\frac{3}{2})_n(\beta+1)_m m!}. \tag{7.12}$$

So we have the expansion

$$\frac{R^{\alpha,\beta,\gamma}_{n,n}(\xi,\eta)}{R^{\alpha,\beta,\gamma}_{n,n}(1,0)}=\sum_{m=0}^{n}\frac{(-n)_m(n+\alpha+\beta+\gamma+\frac{3}{2})_m}{(\beta+1)_m m!}(1-\xi)^m R^{(\alpha,\beta)}_m(1+2(1-\xi)^{-1}\eta). \tag{7.13}$$

In the case $\gamma=-\frac{1}{2}$ formulas (7.11) and (7.13) together are equivalent with Theorem 2.1. For $\xi=1$ we obtain

$$\begin{aligned}\frac{R^{\alpha,\beta,\gamma}_{n,n}(1,\eta)}{R^{\alpha,\beta,\gamma}_{n,n}(1,0)}&=\sum_{m=0}^{n}\frac{(-n)_m(n+\alpha+\beta+\gamma+\frac{3}{2})_m(m+\alpha+\beta+1)_m}{(\beta+1)_m(\alpha+1)_m m!}\eta^m\\&={}_4F_3\left(\begin{matrix}-n,\, n+\alpha+\beta+\gamma+\frac{3}{2},\, \frac{1}{2}(\alpha+\beta+1),\, \frac{1}{2}(\alpha+\beta+2);\\ \alpha+1,\, \beta+1,\, \alpha+\beta+1;\end{matrix}\; 4\eta\right).\end{aligned} \tag{7.14}$$

This formula generalizes (2.26).

Now we can prove the following interesting theorem which connects $R^{\alpha,\beta,\gamma}_{n,n}(\xi,\eta)$ with Appell's function F_4 (cf. (2.28)).

THEOREM 7.5. *We have*

$$\begin{aligned}\frac{R^{\alpha,\beta,\gamma}_{n,n}(\xi,\eta)}{R^{\alpha,\beta,\gamma}_{n,n}(1,0)}&=F_4(-n,n+\alpha+\beta+\gamma+\tfrac{3}{2};\alpha+1,\beta+1;\eta,1-\xi+\eta)\\&=\sum_{i+j\leqq n}\frac{(-n)_{i+j}(n+\alpha+\beta+\gamma+\frac{3}{2})_{i+j}}{(\alpha+1)_i(\beta+1)_j i!j!}\eta^i(1-\xi+\eta)^j.\end{aligned} \tag{7.15}$$

Proof. It follows from (2.5) that

$$(1-\xi)^m R^{(\alpha,\beta)}_m(1+2(1-\xi)^{-1}\eta)=(\beta+1)_m m!\sum_{i+j=m}\frac{\eta^i(1-\xi+\eta)^j}{(\alpha+1)_i(\beta+1)_j i!j!}.$$

Substitution of this formula in (7.13) proves the theorem. □

This theorem generalizes (2.29).

Next we will prove that the polynomials

$$Z^{\gamma}_{n,k}(\xi,\eta)\quad\text{and}\quad(1-\xi)^n R^{(\alpha,\beta)}_k(1+2(1-\xi)^{-1}\eta)$$

can be obtained as limit cases of $R^{\alpha,\beta,\gamma}_{n,k}(\xi,\eta)$ for $\beta\to\infty$, $\gamma\to\infty$, respectively. Thus, because of (2.21), the expansions (4.10) and (7.5) are quite similar to the expansion (2.4) for the Jacobi polynomials. First we note

LEMMA 7.6. *We have*

$$\lim_{\beta\to\infty} \frac{R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta)}{R_{n,0}^{\alpha,\beta,\gamma}(2,1)} = \frac{(\gamma+\frac{1}{2})_n}{(2\gamma+1)_n} Z_{n,0}^{\gamma}(\xi,\eta), \tag{7.16}$$

$$\lim_{\gamma\to\infty} R_{n,n}^{\alpha,\beta,\gamma}(\xi,\eta) = (1-\xi)^n R_n^{(\alpha,\beta)}(1+2(1-\xi)^{-1}\eta). \tag{7.17}$$

Proof. Use (6.2), (3.17), (7.13) and (6.16). □

THEOREM 7.7. *We have*

$$\lim_{\beta\to\infty} \frac{R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)}{R_{n,k}^{\alpha,\beta,\gamma}(2,1)} = \frac{(\gamma+\frac{1}{2})_{n-k}}{(2\gamma+1)_{n-k}} Z_{n,k}^{\gamma}(\xi,\eta), \tag{7.18}$$

$$\lim_{\gamma\to\infty} R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta) = (1-\xi)^n R_k^{(\alpha,\beta)}(1+2(1-\xi)^{-1}\eta). \tag{7.19}$$

Proof. In order to prove (7.18) use complete induction with respect to k. For $k=0$, (7.18) becomes (7.16). From (3.17) and (3.14) we have

$$\frac{R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)}{R_{n,k}^{\alpha,\beta,\gamma}(2,1)} = \frac{1}{4(\beta+1)(\beta+\gamma+\frac{3}{2})} D_+^{\alpha,\beta,\gamma} \frac{R_{n-1,k-1}^{\alpha+1,\beta+1,\gamma}(\xi,\eta)}{R_{n-1,k-1}^{\alpha+1,\beta+1,\gamma}(2,1)}.$$

For $\beta\to\infty$, (3.11) together with the induction hypothesis gives

$$\begin{aligned}\lim_{\beta\to\infty} \frac{R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)}{R_{n,k}^{\alpha,\beta,\gamma}(2,1)} &= \eta \lim_{\beta\to\infty} \frac{R_{n-1,k-1}^{\alpha+1,\beta+1,\gamma}(\xi,\eta)}{R_{n-1,k-1}^{\alpha+1,\beta+1,\gamma}(2,1)} \\ &= \frac{(\gamma+\frac{1}{2})_{n-k}}{(2\gamma+1)_{n-k}} Z_{n,k}^{\gamma}(\xi,\eta).\end{aligned}$$

In order to prove (7.19) we use complete induction with respect to $n-k$, the case $n-k=0$ being clear from (7.17). It follows from (3.16), (3.12) together with the induction hypothesis that

$$\begin{aligned}\lim_{\gamma\to\infty} R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta) &= (1-\xi) \lim_{\gamma\to\infty} R_{n-1,k}^{\alpha,\beta,\gamma+1}(\xi,\eta) \\ &= (1-\xi)^n R_k^{(\alpha,\beta)}(1+2(1-\xi)^{-1}\eta). \quad \square\end{aligned}$$

Remark 7.8. Let us consider the recurrence relations

$$(1-\xi) R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta) = \sum_{m,l} b_{m,l} R_{m,l}^{\alpha,\beta,\gamma}(\xi,\eta), \tag{7.20}$$

$$\eta R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta) = \sum_{m,l} c_{m,l} R_{m,l}^{\alpha,\beta,\gamma}(\xi,\eta). \tag{7.21}$$

By the use of the expansion (7.5), Theorem 7.4 and the orthogonality of the polynomials $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$ it is directly proved that the coefficients $b_{m,l}$ in (7.20) are nonzero only if $(m,l)\in\{(n+1,k), (n,k+1), (n,k), (n,k-1), (n-1,k)\}$ (cf. Fig. 5). Similarly, the coefficients $c_{m,l}$ in (7.21) are nonzero only if $(m,l)\in\{(n+1,k+1), (n+1,k), (n+1,k-1), (n,k+1), (n,k), (n,k-1), (n-1,k+1), (n-1,k), (n-1,k-1)\}$ (cf. Fig. 6). This can be proved by use of the expansion (4.10) and Theorem 4.3. The proofs sketched here are much shorter than those given in Sprinkhuizen [28, § 9].

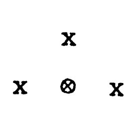

FIG. 5

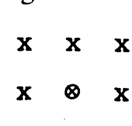

FIG. 6

8. Connection coefficients. In this section we shall consider the connection coefficients in the formula

$$R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)=\sum_{m,l} c_{n,k;m,l} R_{m,l}^{a,b,c}(\xi,\eta). \tag{8.1}$$

It will turn out that for $k=0$ or $k=n$ these coefficients coincide with certain connection coefficients for Jacobi polynomials. If $k=n$ and $(a,b,c)=(\alpha,\beta,-\frac{1}{2})$, or $k=0$ and $(a,b,c)=(-\frac{1}{2},-\frac{1}{2},\gamma)$ then we obtain explicit expressions of $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$, $k=0$ or m, as double Jacobi series. In these cases there will follow important inequalities. We conclude this section by deriving integral representations for $R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)$, $k=0$ or n, in terms of Jacobi polynomials.

First note the following corollary of Theorem 3.2.

LEMMA 8.1. *The coefficients* $c_{n,k;m,l}$ *in* (8.1) *are nonzero only if* $m\leqq n$ *and* $m+l\leqq n+k$ (cf. Fig. 2 in § 3).

THEOREM 8.2.

(a) *The coefficients* $c_{n,k;m,l}$ *in the formula*

$$R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)=\sum_{m,l} c_{n,k;m,l} R_{m,l}^{a,b,\gamma}(\xi,\eta)$$

are nonzero only if $m\leqq n$ *and* $l\leqq k$ (cf. Fig. 3 in § 4).

(b) *The coefficients* $c_{n,k;m,l}$ *in the formula*

$$R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)=\sum_{m,l} c_{n,k;m,l} R_{m,l}^{\alpha,\beta,c}(\xi,\eta)$$

are nonzero only if $m-l\leqq n-k$ *and* $m+l\leqq n+k$ (cf. Fig. 4 in § 7).

Proof. In both cases we can first use Lemma 8.1. Part (a) of the theorem follows by $(k+1)$-fold application of the operator D_-^{γ} to both sides of the formula and by use of (3.13). Similarly, in view of (3.15), part (b) of the theorem is proved by $(n-k+1)$-fold application of $E_-^{\alpha,\beta}$ to both sides of the formula. □

Let the coefficients $g_{n;k}^{\alpha,\beta;a,b}$ be defined by (2.36).

THEOREM 8.3. *There are expansions*

$$R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta)=\sum_{m=0}^{n} g_{n;m}^{\alpha+\gamma+1/2,\beta+\gamma+1/2;a+\gamma+1/2,b+\gamma+1/2} R_{m,0}^{a,b,\gamma}(\xi,\eta), \tag{8.2}$$

$$R_{n,n}^{\alpha,\beta,\gamma}(\xi,\eta)=\sum_{m=0}^{n} g_{n;m}^{\alpha+\gamma+1/2,\beta;\alpha+c+1/2,\beta} R_{m,m}^{\alpha,\beta,c}(\xi,\eta). \tag{8.3}$$

Proof. It follows from Theorem 8.2(a) that $R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta)$ can be expanded in terms of $R_{m,0}^{a,b,\gamma}(\xi,\eta)$, $m=0,1,\cdots,n$. Now restrict to $\eta=\frac{1}{4}\xi^2$ and apply (5.6) and (2.36). This proves (8.2). Similarly, for the proof of (8.3) use Theorem 8.2(b), restrict to $\eta=0$ and apply (5.5) and (2.36). □

The coefficients in (8.3) are positive if $\gamma>c$ (cf. (2.38)). See Theorem 2.3 for the cases that the coefficients in (8.2) are positive.

THEOREM 8.4.

(a) *If*

$$R_n^{(\alpha+\gamma+1/2,\beta+\gamma+1/2)}(x)=\sum_{m=0}^{n} c_{n;m} R_m^{(\gamma,\gamma)}(x)$$

then

$$R_{n,0}^{\alpha,\beta,\gamma}(1-xy,\tfrac{1}{4}(x-y)^2)=\sum_{m=0}^{n} c_{n;m} R_m^{(\gamma,\gamma)}(x) R_m^{(\gamma,\gamma)}(y)$$

and

$$c_{n;m}=g_{n;m}^{\alpha+\gamma+1/2,\beta+\gamma+1/2;\gamma,\gamma}$$
$$=\frac{(n+\alpha+\beta+2\gamma+2)_m(\gamma+1)_m n!}{(m+2\gamma+1)_m(\alpha+\gamma+\frac{3}{2})_m(n-m)!m!}$$
$$\cdot {}_3F_2\left(\begin{matrix}-n+m, n+m+\alpha+\beta+2\gamma+2, m+\gamma+1;\\ 2m+2\gamma+2, m+\alpha+\gamma+\frac{3}{2};\end{matrix}\ 1\right).$$

If either $\alpha>\beta$ and $\alpha+\beta\geqq-1$ or $\alpha=\beta>-\frac{1}{2}$ and $n-m$ is even then in the above formulas $c_{n;m}>0$.

(b) *If*

$$R_n^{(\alpha+\gamma+1/2,\beta)}(x)=\sum_{m=0}^{n} c_{n;m}R_m^{(\alpha,\beta)}(x)$$

then

$$R_{n,n}^{\alpha,\beta,\gamma}(1-\tfrac{1}{2}(x+y),\tfrac{1}{4}(1-x)(1-y))=\sum_{m=0}^{n} c_{n;m}R_m^{(\alpha,\beta)}(x)R_m^{(\alpha,\beta)}(y)$$

and

$$c_{n;m}=g_{n;m}^{\alpha+\gamma+1/2,\beta;\alpha,\beta}$$
$$=\frac{n!(\beta+1)_n(\gamma+\frac{1}{2})_{n-m}(n+\alpha+\beta+\gamma+\frac{3}{2})_m}{(\alpha+\gamma+\frac{3}{2})_n(\alpha+\beta+2)_n(n-m)!(n+\alpha+\beta+2)_m}\omega_m^{(\alpha,\beta)}.$$

If $\gamma>-\frac{1}{2}$ then in the above formulas $c_{n;m}>0$.

Proof. Part (a) of the theorem follows from (8.2) and (3.21). The coefficients are given in (2.39) and Theorem 2.3 implies the positivity result. For part (b) use (8.3), (3.4) and (2.37). □

Theorem 8.4(a) gives an explicit expression for $R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta)$ and it shows that $R_{n,0}^{\alpha,\beta,\gamma}(1-xy,\frac{1}{4}(x-y)^2)$ is the generalized translate of the Jacobi polynomial $R_n^{(\alpha+\gamma+1/2,\beta+\gamma+1/2)}(x)$ expressed as a Gegenbauer series of order (γ,γ) (see Askey [2, Lecture 2] for the definition of generalized translates).

Similarly, Theorem 8.4(b) gives an explicit expression for $R_{n,n}^{\alpha,\beta,\gamma}(\xi,\eta)$ and it shows that $R_{n,n}^{\alpha,\beta,\gamma}(1-\frac{1}{2}(x+y),\frac{1}{4}(1-x)(1-y))$ is the generalized translate of the Jacobi polynomial $R_n^{(\alpha+\gamma+1/2,\beta)}(x)$ expressed as a Jacobi series of order (α,β). Hence we also have a new expression for the generalized translate of the Jacobi polynomial kernel (cf. Bavinck [5, § 5.8], [6]). This kernel gives a summation method for Fourier–Jacobi expansions. If $\gamma\to\infty$ in Theorem 8.4(b) then, using (7.17), we obtain Bateman's bilinear sum, which can be interpreted as the De la Vallée–Poussin kernel (cf. Askey [1]).

For $\gamma=\frac{1}{2}$, Theorem 8.4(b) implies (3.6), and thus the Christoffel–Darboux formula (2.20) for Jacobi polynomials.

COROLLARY 8.5.

(a) *Let $\alpha\geqq\beta$, $\alpha+\beta\geqq-1$, $(\alpha,\beta)\neq(-\frac{1}{2},-\frac{1}{2})$. Then $R_{n,0}^{\alpha,\beta,\gamma}(\xi,0)>0$ for $\xi\in[0,1]$ except if $\alpha=\beta$, n is odd, $\xi=1$. If $\alpha\geqq\beta$, $\alpha+\beta\geqq-1$, $\gamma\geqq-\frac{1}{2}$ then*

$$|R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta)|\leqq 1\quad on\ \bar{\Omega}.$$

(b) *Let $\alpha\leqq\beta$, $\alpha+\beta\geqq-1$, $(\alpha,\beta)\neq(-\frac{1}{2},-\frac{1}{2})$. Then $(-1)^nR_{n,0}^{\alpha,\beta,\gamma}(\xi,\xi-1)>0$ for $\xi\in[1,2]$ except if $\alpha=\beta$, n is odd, $\xi=1$. If $\alpha\leqq\beta$, $\alpha+\beta\geqq-1$, $\gamma\geqq-\frac{1}{2}$ then*

$$|R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta)|\leqq|R_{n,0}^{\alpha,\beta,\gamma}(2,1)|\quad on\ \bar{\Omega}.$$

(c) *If $\gamma>-\frac{1}{2}$ then $R_{n,n}^{\alpha,\beta,\gamma}(\xi,\frac{1}{4}\xi^2)>0$ for $\xi\in[0,2]$. If $\gamma\geqq\frac{1}{2}$,* $\max(\alpha,\beta)\geqq-\frac{1}{2}$, $(\xi,\eta)\in$

$\bar{\Omega}$ *then*

$$|R_{n,n}^{\alpha,\beta,\gamma}(\xi,\eta)| \leqq \begin{cases} 1 & \text{if } \alpha \geqq \beta, \\ |R_{n,n}^{\alpha,\beta,\gamma}(2,1)| & \text{if } \alpha \leqq \beta. \end{cases}$$

Proof. (a) It follows from Theorem 8.4(a) that

$$R_{n,0}^{\alpha,\beta,\gamma}(1-x^2,0) = \sum_{m=0}^{n} c_{n;m}(R_m^{(\gamma,\gamma)}(x))^2.$$

If $\alpha > \beta$, $\alpha+\beta \geqq -1$ then $c_{n;n}$ and $c_{n;n-1}$ are both positive. By Szegö [29, Thm. 3.3.3] $R_n^{(\gamma,\gamma)}(x)$ and $R_{n-1}^{(\gamma,\gamma)}(x)$ cannot have common zeros. Hence $R_{n,0}^{\alpha,\beta,\gamma}(1-x^2,0)>0$ for $0 \leqq x \leqq 1$. If $\alpha=\beta>-\frac{1}{2}$ then $c_{n;n}$ and $c_{n;n-2}$ are both positive. The positivity results again from [29, Thm. 3.3.3] together with (2.9) and (2.10). The second statement follows from Theorem 8.4(a) and the fact that $|R_m^{(\gamma,\gamma)}(x)| \leqq 1$ if $\gamma \geqq -\frac{1}{2}$ and $-1 \leqq x \leqq 1$; cf. Theorem 2.2.

(b) Use part (a) of the corollary together with (3.18).

(c) It follows from Theorem 8.4(b) that

$$R_{n,n}^{\alpha,\beta,\gamma}(1-x, \tfrac{1}{4}(1-x)^2) = \sum_{m=0}^{n} c_{n;m}(R_m^{(\alpha,\beta)}(x))^2.$$

A similar argument as in the proof of (a) gives the positivity result. The second statement follows from Theorem 8.4(b) together with the inequalities for Jacobi polynomials (cf. Theorem 2.2). □

The above corollary confirms part of the hypothesis that for $\alpha \geqq \beta \geqq -\frac{1}{2}$, $\gamma \geqq -\frac{1}{2}$ the inequality

$$|R_{n,k}^{\alpha,\beta,\gamma}(\xi,\eta)| \leqq 1$$

is valid on $\bar{\Omega}$; cf. Sprinkhuizen [28, § 7].

Let us conclude this paper by deriving integral representations for $R_{n,0}^{\alpha,\beta,\gamma}(\xi,\eta)$ and $R_{n,n}^{\alpha,\beta,\gamma}(\xi,\eta)$. Combination of Theorem 8.4(a) and formula (2.33) gives

$$\begin{aligned} R_{n,0}^{\alpha,\beta,\gamma}(1-xy, \tfrac{1}{4}(x-y)^2) &= \frac{\Gamma(\gamma+1)}{\pi^{1/2}\Gamma(\gamma+\frac{1}{2})} \\ &\cdot \int_{-1}^{1} R_n^{(\alpha+\gamma+1/2,\beta+\gamma+1/2)}(xy+(1-x^2)^{1/2}(1-y^2)^{1/2}t)(1-t^2)^{\gamma-1/2}\,dt, \qquad \gamma > -\tfrac{1}{2}, \end{aligned} \tag{8.4}$$

which can be considered as a generalization of the product formula (2.33).

Similarly, Theorem 8.4(b) and formula (2.35) imply the following generalization of the product formula (2.35):

$$\begin{aligned} &R_{n,n}^{\alpha,\beta,\gamma}(1-\tfrac{1}{2}(x+y), \tfrac{1}{4}(1-x)(1-y)) \\ &= \frac{2\Gamma(\alpha+1)}{\pi^{1/2}\Gamma(\alpha-\beta)\Gamma(\beta+\frac{1}{2})} \\ &\cdot \int_0^1 \int_0^{\pi} R_n^{(\alpha+\gamma+1/2,\beta)}(\tfrac{1}{2}(1+x)(1+y)+\tfrac{1}{2}(1-x)(1-y)r^2 \\ &\qquad + (1-x^2)^{1/2}(1-y^2)^{1/2} r\cos\phi - 1) \\ &\qquad \cdot (1-r^2)^{\alpha-\beta-1} r^{2\beta+1} (\sin\phi)^{2\beta}\,dr\,d\phi, \qquad \alpha > \beta > -\tfrac{1}{2}. \end{aligned} \tag{8.5}$$

Both in (8.4) and (8.5) the left hand sides can be considered as the first term of an orthogonal expansion of the integrand. The full orthogonal expansion (a generalized addition formula) can be obtained by means of the techniques described in Koornwinder [24], i.e. by using integration by parts and differential recurrence relations for Jacobi polynomials. In particular, from (8.4) we get

$$\begin{aligned} &R_n^{(\alpha+\gamma+1/2,\beta+\gamma+1/2)}(xy+(1-x^2)^{1/2}(1-y^2)^{1/2}t) \\ &= \sum_{k=0}^{n} \frac{(-1)^k(-n)_k(n+\alpha+\beta+2\gamma+2)_k}{2^{2k}(\alpha+\gamma+\frac{3}{2})_k(\gamma+1)_k}(1-x^2)^{k/2}(1-y^2)^{k/2} \\ &\quad \cdot R_{n-k,0}^{\alpha,\beta,\gamma+k}(1-xy,\tfrac{1}{4}(x-y)^2)\omega_k^{(\gamma-1/2,\gamma-1/2)}R_k^{(\gamma-1/2,\gamma-1/2)}(t), \end{aligned} \tag{8.6}$$

which is a generalization of the addition formula (2.34) for Gegenbauer polynomials. See Manocha [26] and Carlson [9, § 3] for related generalizations of this addition formula.

Notes added in proof.

1. The hypothesis quoted in Remark 6.12 has been proved by the second author [30]. It is shown there that (6.19) is valid if $\alpha \geqq -\frac{1}{2}$, $\alpha \geqq \beta$, $\alpha+\beta+1 \geqq 0$ and $\gamma \geqq -\frac{1}{2}$.

2. See Koornwinder [31] for a proof that the polynomials under consideration of order $(\frac{1}{2}(d-5), -\frac{1}{2}, 0)$ can be interpreted as spherical functions on the Grassmann manifold $O(d)/O(2)\times O(d-2)$ (cf. the remarks after (4.15)).

3. Our Corollary 6.6 was independently obtained by K. Ringhofer, Universität Osnabrück, B.R.D. (private communication). He also showed that these orthogonal polynomials naturally arise in the decomposition of the Kronecker product of two most degenerate discrete representations of the conformal group $SO_0(4, 2)$ (still unpublished work).

REFERENCES

[1] R. ASKEY, *Jacobi polynomials, I. New proofs of Koornwinder's Laplace type integral representation and Bateman's bilinear sum*, this Journal, 5 (1974), pp. 119–124.

[2] ———, *Orthogonal Polynomials and Special Functions*, Regional Conference Series in Applied Mathematics, no. 21, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1975.

[3] R. ASKEY AND G. GASPER, *Jacobi polynomial expansions of Jacobi polynomials with non-negative coefficients*, Proc. Cambridge Philos. Soc., 70 (1971), pp. 243–255.

[4] H. BATEMAN, *Partial differential equations of mathematical physics*, Cambridge University Press, Cambridge, England, 1932.

[5] H. BAVINCK, *Jacobi Series and Approximation*, Mathematical Centre Tracts, no. 39, Mathematisch Centrum, Amsterdam, 1972.

[6] ———, *Convolution operators for Fourier-Jacobi expansions*, Linear Operators and Approximation, P. L. Butzer, J.-P. Kahane and B. Sz.-Nagy, eds., ISNM vol. 20, Birkhäuser-Verlag, Basel, Switzerland, 1972, pp. 371–380.

[7] J. L. BURCHNALL AND T. W. CHAUNDY, *Expansions of Appell's double hypergeometric functions, II*, Quarterly J. Math. Oxford. Ser., 12 (1941), pp. 112–128.

[8] B. C. CARLSON, *Appell functions and multiple averages*, this Journal, 2 (1971), pp. 420–430.

[9] ———, *Quadratic transformations of Appell functions*, this Journal, 7 (1976), pp. 291–304.

[10] A. G. CONSTANTINE, *Some non-central distribution problems in multivariate analysis*, Ann. Math. Statist., 34 (1963), pp. 1270–1285.

[11] J. DOUGALL, *A theorem of Sonine in Bessel functions, with two extensions to spherical harmonics*, Proc. Edinburgh Math. Soc., 25 (1907), pp. 114–132.

[12] A. ERDÉLYI, W. MAGNUS, F. OBERHETTINGER AND F. G. TRICOMI, *Higher Transcendental Functions*, vol. I, McGraw-Hill, New York, 1953.

[13] ———, *Higher Transcendental Functions*, vol. II, McGraw-Hill, New York, 1953.

[14] E. FELDHEIM, *Contributions à la theorie des polynomes de Jacobi*, Mat. Fiz. Lapok, 48 (1941), pp. 453–504. (Hungarian, French summary.)

[15] G. GASPER, *Linearization of the product of Jacobi polynomials, I*, Canad. J. Math., 22 (1970), pp. 171–175.

[16] ———, *Linearization of the product of Jacobi polynomials, II*, Ibid., 22 (1970), pp. 582–593.

[17] C. S. HERZ, *Bessel functions of matrix argument*, Ann. of Math., 61 (1955), pp. 474–523.

[18] A. T. JAMES, *Zonal polynomials of the real positive definite symmetric matrices*, Ibid., 74 (1961), pp. 456–469.

[19] ———, *Calculation of zonal polynomial coefficients by use of the Laplace–Beltrami operator*, Ann. Math. Statist., 39 (1968), pp. 1711–1718.

[20] A. T. JAMES AND A. G. CONSTANTINE, *Generalized Jacobi polynomials as spherical functions of the Grassmann manifold*, Proc. London Math. Soc. (3), 29 (1974), pp. 174–192.

[21] T. H. KOORNWINDER, *Orthogonal polynomials in two variables which are eigenfunctions of two algebraically independent partial differential operators, I, II*, Nederl. Akad. Wetensch. Proc. Ser. A, 77 = Indag. Math., 36 (1974), pp. 59–66.

[22] ———, *Jacobi polynomials II. An analytic proof of the product formula*, this Journal, 5 (1974), pp. 125–137.

[23] ———, *The addition formula for Jacobi polynomials and the theory of orthogonal polynomials in two variables, a survey*, Math. Centrum Amsterdam Rep. TW 145, 1974.

[24] ———, *Jacobi polynomials, III. An analytic proof of the addition formula*, this Journal, 6 (1975), pp. 533–543.

[25] ———, *Two-variable analogues of the classical orthogonal polynomials*, Theory and Application of Special Functions, R. Askey, ed., Academic Press, New York, 1975, pp. 435–495.

[26] H. L. MANOCHA, *Some formulae involving Appell's function F_4*, Publ. Inst. Math. (Beograd), 9 (23) (1969), pp. 153–156.

[27] L. J. SLATER, *Generalized hypergeometric functions*, Cambridge University Press, Cambridge, England, 1966.

[28] I. G. SPRINKHUIZEN-KUYPER, *Orthogonal polynomials in two variables. A further analysis of the polynomials orthogonal over a region bounded by two lines and a parabola*, this Journal, 7 (1976), pp. 501–518.

[29] G. SZEGÖ, *Orthogonal Polynomials*, Colloquium Publications, vol. 23, 3rd ed., American Mathematical Society, Providence, RI, 1967.

[30] I. G. SPRINKHUIZEN-KUYPER, *A Jacobi series expansion with nonnegative coefficients related to a special class of orthogonal polynomials in two variables*, Rep. TW 163, Math. Centrum, Amsterdam, 1976.

[31] T. H. KOORNWINDER, *Harmonics and spherical functions on Grassmann manifolds of rank two and two-variable analogues of Jacobi polynomials*, Constructive Theory of Functions of Several Variables, W. Schempp and K. Zeller, eds., Lecture Notes in Mathematics 571, Springer-Verlag, Berlin, 1977, pp. 141–154.

SIAM J. MATH. ANAL.
Vol. 9, No. 3, June 1978

Copyright © 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0903-0004$1.00/0

WIRTINGER'S INEQUALITY*

C. A. SWANSON†

Abstract. General forms of Wirtinger-type inequalities are proved in both one and n dimensions. Since singular endpoints and unbounded intervals are allowed, a large class of new one-dimensional results are generated as well as previously known results. In the (usual) case that the admissible functions are identically zero on the boundary ∂G of a bounded domain G in E^n, the sharp form of Wirtinger's inequality in G is proved without any regularity hypotheses on ∂G. If the admissible functions are not so restricted, the companion inequality is proved for domains with $\mathbf{C}^2$ boundaries.

1. Introduction. Generalizations of the classical one-dimensional (quadratic) Wirtinger inequality are proved and the results are extended to bounded domains in n-dimensional Euclidean space. The one-dimensional theorems include all previously given verisions of quadratic Wirtinger-type inequalities, by either specialization or multiple application in subintervals. The proof of the basic Theorem 1, allowing singular endpoints and/or unbounded intervals, is *much* easier and more direct than that given for special cases of it by Beesack [1], Diaz and Metcalf [2], Fan, Taussky and Todd [3], Hardy, Littlewood and Pólya [4], and others. The class of admissible functions (to which our inequality applies) consists of absolutely continuous functions in appropriate weighted $\mathbf{L}^2$-spaces. Restriction of this natural class to $\mathbf{C}^1(\bar{I})$ yields classical Wirtinger inequalities on closed bounded intervals I.

For analogous inequalities in bounded n-dimensional domains G, the admissibility class is taken to be $\mathfrak{D}_0 = \mathbf{C}(\bar{G}) \cap \mathbf{H}_2^1(G)$, where $\mathbf{H}_2^1(G)$ denotes a standard Sobolev space (described in § 3). This is a natural extension of the class used in Theorem 2 in the case of bounded intervals I. The one-dimensional method is not easily extended since limits of quotients $u^2(x)/v(x)$ as x approaches the boundary cannot be found directly by L'Hôpital's rule. The results in Theorems 5 and 6 are obtained, according as $\mathfrak{D}_0$ is restricted to functions vanishing identically on ∂G or not so restricted, by two independent methods. In the first of these (Theorem 5), we use the fact that a function $u \in \mathfrak{D}_0$ with $u \equiv 0$ on ∂G can be approximated arbitrarily closely by functions of class $\mathbf{C}_0^\infty(G)$, and this requires no special boundary regularity hypothesis (Lemma 4). In Theorem 6, we use the Hopf maximum principle, and accordingly must require that ∂G be of class $\mathbf{C}^2$.

The question of how much boundary regularity is needed seems to have been largely ignored in the literature, or else too much or too little regularity has been presupposed. The proof offered by Wong [9] of an inequality of type (23) or (25) contains an error, described following Theorem 6. Weak forms of (23), (25), not characterizing the case of equality, are essentially contained in the author's earlier papers [6], [7], [8].

Wirtinger-type inequalities have a close connection with the Euler–Lagrange differential equations associated with variational problems of mathematical physics, in particular the isoperimetric problems. The well-known characterization of the smallest eigenvalue of a Sturm–Liouville system as the minimum of the corresponding Rayleigh quotient implies weak Wirtinger inequalities, but this approach does not give insight into the sharp results contained herein.

---

* Received by the editors June 4, 1976.

† Department of Mathematics, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5.

2. One-dimensional Wirtinger-type inequalities. Let L be the differential operator defined by

$$Lv=(Av')'+Bv \tag{1}$$

in an open interval $I=(a,b)$, which is permitted to be unbounded, where A and B are continuous functions in I and $A(x)>0$ in I. The domain $\mathfrak{D}_L$ of L is defined to be the set of all real-valued functions v in I such that all derivatives involved in (1) exist and are continuous at each point in I. We shall consider solutions $v\in\mathfrak{D}_L$ of the differential inequality

$$-Lv\geqq\lambda_0 Cv \tag{2}$$

in I, where λ_0 is a real number and C is a positive continuous function in I.

The notation $\mathbf{L}_A^2(I)$ will be used for the set of all real-valued measurable functions u in I such that Au^2 is Lebesgue integrable in I. Also $\mathbf{AbC}(I)$ will denote the set of all real-valued functions which are absolutely continuous on every closed subinterval of I. For a positive solution $v\in\mathfrak{D}_L$ of (2), we consider functions $u\in\mathbf{AbC}(I)$ such that the limits below exist and are finite:

$$S_1(u,v)=\lim_{x\to a+}\frac{A(x)u^2(x)v'(x)}{v(x)},\qquad S_2(u,v)=\lim_{x\to b-}\frac{A(x)u^2(x)v'(x)}{v(x)}. \tag{3}$$

THEOREM 1. *Let $v\in\mathfrak{D}_L$ be a positive solution of* (2) *in I for some real number λ_0. Then every $u\in\mathbf{AbC}(I)$ such that $u\in\mathbf{L}_B^2(I)\cap\mathbf{L}_C^2(I)$, $u'\in\mathbf{L}_A^2(I)$, and the limits* (3) *exist and are finite satisfies the inequality*

$$\int_I(B+\lambda_0C)u^2\,dx\leqq\int_I A(u')^2\,dx+S_1(u,v)-S_2(u,v). \tag{4}$$

Furthermore, equality holds if and only if $u(x)$ is a constant multiple of $v(x)$ throughout I.

Proof. The differential identity

$$Av^2\left[\left(\frac{u}{v}\right)'\right]^2+\left[\frac{Au^2v'}{v}\right]'=A(u')^2-Bu^2+\frac{u^2}{v}Lv$$

can be integrated over a subinterval (y,z) of I, $a<y<z<b$, to yield, since u is absolutely continuous,

$$\int_y^z[A(u')^2-(B+\lambda_0C)u^2]\geqq\left[\frac{Au^2v'}{v}\right]_y^z, \tag{5}$$

equality holding if and only if $(u/v)'=0$ a.e. in I, i.e. $u(x)=(\text{const.})v(x)$ identically in I. Taking limits as $y\to a+$ and $z\to b-$, we obtain the conclusion (4).

As one specialization of Theorem 1, consider the case that $I=(a,b)$ is bounded, thd continuity of A, B, and C extends to $[a,b]$, and v is a positive eigenfunction corresponding to the smallest eigenvalue λ_0 of the problem

$$\begin{aligned}&Lv+\lambda_0Cv=0\quad\text{in } I,\\&s_1v(a)-A(a)v'(a)=0,\\&s_2v(b)+A(b)v'(b)=0,\end{aligned} \tag{6}$$

where s_1 and s_2 are constants. (The cases $s_1=\infty$, $s_2=\infty$ by convention correspond to boundary conditions $v(a)=0$, $v(b)=0$, respectively.) For finite s_1 and s_2, conditions

(3) become

$$S_1(u, v) = s_1 u^2(a), \qquad S_2(u, v) = -s_2 u^2(b). \tag{7}$$

In the case of null boundary conditions $v(a)=0$ and/or $v(b)=0$, we need the additional assumptions that $u(a)=0$ and/or $u(b)=0$ to conclude that $S_1(u, v)=0$ and/or $S_2(u, v)=0$. If $u \in \mathbf{C}^1[a, b]$ this follows from L'Hôpital's rule since a solution v of the differential equation (6) can have only simple zeros. However, as the proof of Theorem 5 shows, the additional hypothesis that $u \in \mathbf{C}^1[a, b]$ is unnecessary.

THEOREM 2. *Let $v \in \mathfrak{D}_L$ be a positive-valued eigenfunction in a bounded interval I corresponding to the smallest eigenvalue λ_0 of* (6). *Then every absolutely continuous function u on $[a, b]$ such that $u' \in \mathbf{L}^2(a, b)$ satisfies the Wirtinger-type inequality*

$$\int_a^b (B + \lambda_0 C)u^2 \, dx \leqq \int_a^b A(u')^2 \, dx + s_1 u^2(a) + s_2 u^2(b), \tag{8}$$

equality holding if and only if $u(x) = Kv(x)$ throughout $[a, b]$ for some constant K. (*If $u(a)=0$ and/or $u(b)=0$, it is understood that the corresponding boundary conditions in* (6) *are $v(a)=0$ and/or $v(b)=0$.*)

Beesack's Theorem 1.1 [1] is obtained from Theorem 2 by substituting $A(x) \equiv 1$, $\lambda_0 = 0$, $s_2 \leqq 0$, $s_1 = \infty$ (i.e. $v(a)=0$ in (6)), $u(a)=0$, and $a=0$. Then (8) implies the inequality

$$\int_a^b (u'(x))^2 \, dx \geqq \int_a^b B(x) u^2(x) \, dx, \tag{9}$$

with equality holding if and only if $u(x) = (\text{const.})v(x)$ (*and* if and only if $u(x) \equiv 0$ in the case $s_2 < 0$). Beesack's theorem 1.1* replaces the hypothesis $u(a)=0$ by the "orthogonality" conditions $\int_a^b B(x)\, dx \geqq 0$ and $u(a) \int_a^b B(x) u(x)\, dx \leqq 0$. This follows immediately upon application of (9) to $u_1(x) = u(x) - u(a)$. Theorem 1.2 of Beesack allows $v(x)$ to have a simple interior zero $x_0 \in (a, b)$. The proof of Theorems 1 and 2 goes through if $u(x)$ is replaced by $u_1(x) = u(x) - u(x_0)$, since we need merely to split the integral in (5) into integrals over subintervals (y, x_0), (x_0, z) and use L'Hôpital's rule to conclude that

$$\lim_{x \to x_0} \frac{A(x) u_1^2(x) v'(x)}{v(x)} = 0. \tag{10}$$

Other classical and modern versions of Wirtinger's inequality all follow from Theorem 2 by specialization, or modification by the addition of hypotheses, or multiple application on subintervals. The following corollary is the central theorem used by Diaz and Metcalf [2], also given in modified form by Hardy, Littlewood and Pólya [4].

COROLLARY 3. *Every real-valued function $w \in \mathbf{C}^1[a, b]$ satisfies the inequality*

$$\int_a^b [w(x) - w(a)]^2 \, dx \leqq 4\left(\frac{b-a}{\pi}\right)^2 \int_a^b [w'(x)]^2 \, dx, \tag{11}$$

equality holding if and only if

$$w(x) = w(a) + K \sin\left(\frac{\pi(x-a)}{2(b-a)}\right)$$

identically on $[a, b]$ for some constant K.

This is the special case of (1), (8) in which $A=1$, $B=0$, $C=1$, $\lambda_0=\pi^2/(4(b-a)^2)$, $s_1=\infty$ (i.e. $v(a)=0$), $s_2=0$ (i.e. $v'(b)=0$). We apply Theorem 2 to

$$u(x)=w(x)-w(a), \qquad v(x)=\sin\left[\frac{\pi(x-a)}{2(b-a)}\right]$$

to get (11) immediately.

Several auxiliary theorems of Diaz and Metcalf are obtained by repeated applications of Corollary 3 on subintervals of $[a,b]$. One very interesting result is the classical Wirtinger inequality for periodic $\mathbf{C}^1$ functions u on $[0,\pi]$: If $u(0)=u(2\pi)$ and $\int_0^{2\pi} u(x)\,dx=0$, then

$$\int_0^{2\pi} u^2(x)\,dx \leqq \int_0^{2\pi} [u'(x)]^2\,dx,$$

equality holding if and only if $u(x)=K\sin(x-\delta)$ identically on $[0,2\pi]$ for some real numbers K and δ. This follows by application of Corollary 3 to four subintervals of $(0,2\pi)$ [2].

The following examples illustrate the case of unbounded intervals I.

Example 1. In Theorem 1, suppose that $I=(1,\infty)$, $A(x)=xe^{-x}$, $B(x)=0$, $C(x)=e^{-x}$, and $\lambda_0=1$. If $u\in\mathbf{L}^2_C(1,\infty)$, $u'\in\mathbf{L}^2_A(1,\infty)$, $u(x)=o(\sqrt{x-1})$ as $x\to 1+$, and u is absolutely continuous on $[1,a]$ for arbitrary $a>1$, the conclusion of Theorem 1 is that

$$\int_1^\infty u^2(x)\,e^{-x}\,dx \leqq \int_1^\infty [u'(x)]^2 x\,e^{-x}\,dx,$$

equality iff $u(x)=K(x-1)$ throughout $1\leqq x<\infty$ for some constant K. (The limits (3) are easily seen to be zero for $v(x)=x-1$.)

Example 2. Take $I=(0,\infty)$, $A(x)=C(x)=\exp(-x^2)$, $B(x)=0$, $\lambda_0=2$. If u and u' are in $\mathbf{L}^2_A(0,\infty)$, u is absolutely continuous on $[0,b]$ for arbitrary $b>0$, and $u(x)=o(\sqrt{x})$ $(x\to 0+)$, Theorem 1 gives

$$\int_0^\infty u^2(x)\exp(-x^2)\,dx \leqq \frac{1}{2}\int_0^\infty [u'(x)]^2\exp(-x^2)\,dx,$$

equality holding iff $u(x)=Kx$ in $0\leqq x<\infty$ for some constant K.

In the following example, the differential equation (1) is singular at an endpoint of the (finite) interval I.

Example 3. In Theorem 1, take $I=(0,1)$, $A(x)=1-x^2$, $B(x)=0$, $C(x)=1$, $\lambda_0=2$, $v(x)=x$. If $u\in\mathbf{L}^2(0,1)$, $u'\in\mathbf{L}^2_A(0,1)$, u is absolutely continuous on $[0,1]$, and $u(x)=o(\sqrt{x})(x\to 0+)$, then

$$\int_0^1 u^2(x)\,dx \leqq \frac{1}{2}\int_0^1 (1-x^2)[u'(x)]^2\,dx,$$

equality holding iff $u(x)\equiv Kx$ for some constant K.

It is of course an easy matter to write down similar examples when one or both of the boundary limits (3) are not zero. Combination of several inequalities of type (4) on subintervals $I_1, I_2, \cdots$ of I on which corresponding functions $v_1, v_2, \cdots$ are positive yields various interesting inequalities, for example extensions of the classical Wirtinger result for periodic $\mathbf{C}^1$ functions cited above.

3. Extensions to *n* dimensions. Let G be a bounded domain in n-dimensional Euclidean space E^n with boundary ∂G. Points in E^n are denoted by $x=(x_1,\cdots,x_n)$

and differentiation with respect to x_i is denoted by D_i, $i = 1, \cdots, n$. Let $\mathbf{C}_0^\infty(G)$ denote the set of all infinitely differentiable functions $u: E^n \to E^1$ with compact support contained in G. The space $\mathring{\mathbf{W}}_2^1(G)$ is defined as the completion of $\mathbf{C}_0^\infty(G)$ in the norm $\|\cdot\|$ defined by

$$\|u\|^2 = \int_G \left(|u|^2 + \sum_{i=1}^n |D_i u|^2 \right) dx. \tag{12}$$

As usual $\mathbf{C}^{1*}(G)$ denotes the set of all $u \in \mathbf{C}^1(G)$ with $\|u\|$ finite, and $\mathbf{H}_2^1(G)$ denotes the completion of $\mathbf{C}^{1*}(G)$ in the norm (12). The following lemma [5] is an extension of a well-known result to arbitrary domains G.

LEMMA 4. *If $u \in \mathbf{C}(\bar{G}) \cap \mathbf{H}_2^1(G)$ and if u is identically zero on ∂G, then $u \in \mathring{\mathbf{W}}_2^1(G)$.*

Elliptic differential operators L defined by

$$Lv = \sum_{i,j=1}^n D_i(A_{ij} D_j v) + 2 \sum_{i=1}^n B_i D_i v + Cv \tag{13}$$

will be considered in G, where it is assumed that each $A_{ij} \in \mathbf{C}^1(\bar{G})$, each $B_i \in \mathbf{C}(\bar{G})$, $C \in \mathbf{C}(\bar{G})$, and that the matrix $(A_{ij}(x))$ is positive definite (or semidefinite) in G. The domain $\mathfrak{D}_L$ of L is defined to be $\mathbf{C}^2(G) \cap \mathbf{H}_2^1(G)$.

Let E be a continuous real-valued function on $\bar{G}$ such that the quadratic form

$$Q[z] = \sum_{i,j=1}^n A_{ij} z_i z_j - 2 z_{n+1} \sum_{i=1}^n B_i z_i + E z_{n+1}^2 \tag{14}$$

is positive definite (or semidefinite). If (A_{ij}) is positive definite, a well known sufficient condition for $Q[z]$ to be positive semidefinite is $\det Q \geqq 0$, where Q is the matrix associated with $Q[z]$. The following notation will be used:

$$F[u] = \int_G \left[\sum_{i,j} A_{ij} D_i u D_j u - 2u \sum_i B_i D_i u + (E - C) u^2 \right] dx, \tag{15}$$

$$\Phi[u, v] = \sum_{i,j} v^2 A_{ij} D_i\left(\frac{u}{v}\right) D_j\left(\frac{u}{v}\right) - 2uv \sum_i B_i D_i\left(\frac{u}{v}\right) + Eu^2, \tag{16}$$

$$H_G[u, v] = \int_G \Phi[u(x), v(x)]\, dx \tag{17}$$

whenever they are well-defined. The domain of the functional F is taken to be

$$\mathfrak{D}_0 = \mathbf{C}(\bar{G}) \cap \mathbf{H}_2^1(G).$$

In analogy with the one-dimensional case, we consider positive-valued solutions $v \in \mathfrak{D}_L$ of the differential inequality

$$-Lv \geqq \lambda_0 C_0 v \tag{18}$$

in G, where L is given by (13), λ_0 is a real number, and C_0 is a continuous positive-valued function on $\bar{G}$. There is no loss of generality in taking $\lambda_0 = 0$ (as in Theorem 5 below) since the function C in (1) can be replaced by $C_1 = C + \lambda_0 C_0$.

THEOREM 5. *Suppose that a continuous real-valued function E on $\bar{G}$ has been chosen so that the quadratic form* (14) *is positive definite* (*or semidefinite*) *throughout a bounded domain G. If there exists a positive solution $v \in \mathfrak{D}_L$ of the differential inequality $-Lv \geqq 0$ throughout G, then every function $u \in \mathfrak{D}_0$ which vanishes identically on ∂G satisfies the generalized Wirtinger inequality $F[u] \geqq 0$. Furthermore, in the case that* (14) *is positive definite, the inequality reduces to an equality if and only if $Lv = 0$ in G and $u(x) = Kv(x)$ for some constant K, and $K = 0$ if $E \neq 0$.*

Proof. Since $u \in \mathbf{C}(\bar{G}) \cap \mathbf{H}_2^1(G)$ and u is identically zero on ∂G, Lemma 4 implies that there exists a sequence of functions $u_m \in \mathbf{C}_0^\infty(G)$, $m = 1, 2, \cdots$, such that $\lim \|u_m - u\| = 0$ $(m \to \infty)$. Since $v \in \mathfrak{D}_L$ and $v > 0$ in G, we can use Picone's identity [7]

$$\Phi[u_m, v] + \sum_{i,j} D_i \left[\frac{u_m^2}{v} A_{ij} D_j v\right] = F_I[u_m] + \frac{u_m^2}{v} Lv \tag{19}$$

where $F_I[u]$ denotes the integrand in (15). Since $\Phi[u, v]$ is positive definite or semidefinite, $-Lv \geqq 0$, and $u_m \equiv 0$ outside a compact subset of G, integration of (19) over G yields the inequality

$$F[u_m] \geqq H_G[u_m, v] \geqq 0. \tag{20}$$

On account of the uniform boundedness of each A_{ij}, B_i, C and E in G, the Cauchy–Schwarz inequality gives the estimate

$$|F[u_m] - F[u]| \leqq k(\|u_m\| + \|u\|)\|u_m - u\| \tag{21}$$

for some constant k independent of m. Since $\|u_m - u\| \to 0$ as $m \to \infty$, it follows from (20) that $F[u] \geqq 0$.

In the case $F[u] = 0$, equation (19) shows that the inequality $-Lv \geqq 0$ must reduce to an equality in G, and (21) shows that $F[u_m] \to 0$ as $m \to \infty$. For an arbitrary domain S with $\bar{S} \subset G$, inequality (20) implies that

$$0 \leqq H_S[u_m, v] \leqq F[u_m]. \tag{22}$$

Since v is uniformly bounded away from zero on $\bar{S}$ and $v \in \mathfrak{D}_L$, the Cauchy–Schwarz inequality can be used to derive the following analogue of (21):

$$|H_S[u_m, v] - H_S[u, v]| \leqq k_1\left(\left\|\frac{u_m}{v}\right\|_S + \left\|\frac{u}{v}\right\|_S\right)\left\|\frac{u_m - u}{v}\right\|_S$$

where the subscript S on the norms indicate that the integration in (12) is over S only. Since $F[u_m] \to 0$ it follows from (22) that $H_S[u, v] = 0$ and hence that each $D_i(u/v)$, $i = 1, \cdots, n$, and Eu are identically zero in S when the form Φ in (16), (17) is positive definite. Since S is arbitrary, this means that $u(x) = Kv(x)$ for some constant K, and $K = 0$ if $E \neq 0$. Conversely, it is easily seen from (19) that $F[u] = 0$ if $Lv = 0$ and $u(x) = Kv(x)$ ($K = 0$ if $E \neq 0$).

If L is a symmetric elliptic operator, i.e. $B_i \equiv 0$ in (13) for $i = 1, \cdots, n$, then E can be taken identically zero in (14)–(17), and the Wirtinger inequality reduces to

$$\int_G Cu^2\, dx \leqq \int_G \sum_{i,j=1}^n A_{ij} D_i u D_j u\, dx, \tag{23}$$

with equality holding in the positive definite case if and only if $u(x) = Kv(x)$ for some constant K.

Theorem 2 also can be extended to n dimensions, but it must be assumed that ∂G is of class $\mathbf{C}^2$, so that the Hopf maximum principle can be applied. (There is no analogue of Lemma 4 available when u fails to be identically zero on ∂G.) We assume also that (A_{ij}) is uniformly positive definite in G.

Let $v \in \mathfrak{D}_L \cap \mathbf{C}^1(\bar{G})$ denote a positive-valued eigenfunction corresponding to the smallest eigenvalue λ_0 of the problem

$$(24)\qquad \begin{aligned} Lv &\equiv \sum_{i,j=1}^{n} D_i(A_{ij}D_jv)+Cv=-\lambda_0 C_0 v \quad \text{in } G,\\ S[v] &\equiv Fv+\sum_{i,j=1}^{n} A_{ij}\nu_i D_j v=0 \quad \text{on } \partial G, \end{aligned}$$

where F is a continuous function on ∂G and (ν_i) denotes the exterior unit normal to ∂G.

THEOREM 6. *Suppose that (A_{ij}) in* (24) *is uniformly positive definite in a bounded domain G with $\mathbf{C}^2$ boundary at every point. Let v be a positive eigenfunction of* (24) *corresponding to the smallest eigenvalue λ_0. Then every function $u \in \mathfrak{D}_0$ satisfies the inequality*

$$(25)\qquad \int_G (C+\lambda_0 C_0)u^2\,dx \leqq \int_G \sum_{i,j=1}^{n} A_{ij}D_iuD_ju\,dx+\int_{\partial G} Fu^2\,ds$$

where s denotes the measure on ∂G. Furthermore, equality in (25) *holds if and only if $u(x)=Kv(x)$ for some constant K.*

Proof. Since $\partial G \in \mathbf{C}^2$ and $v(x)>0$ in G, it follows that $v(x)>0$ on $\bar{G}$, for if $v(x^0)=0$ at a boundary point x^0, then (24) would imply that the transverse derivative

$$Dv \equiv \sum_{i,j=1}^{n} A_{ij}\nu_i D_j v=0$$

at x^0, contradicting the Hopf maximum principle. Thus the identity (19) can be integrated over G to give

$$(26)\qquad \begin{aligned} F[u] &\geqq -\int_G \frac{u^2}{v}Lv\,dx+\int_{\partial G}\frac{u^2}{v}\sum_{i,j}A_{ij}\nu_iD_jv\,ds\\ &=\int_G \lambda_0 C_0 u^2\,dx-\int_{\partial G} Fu^2\,ds. \end{aligned}$$

In view of (15) (in the case $B_i \equiv 0$ and $E \equiv 0$), the inequality (26) is equivalent to (25). Furthermore, by (16) and (19), equality holds in (26) if and only if $D_i(u/v)\equiv 0$ in G for $i=1,\cdots,n$, i.e. if and only if $u(x)=Kv(x)$ for some constant K.

Wong [9] has stated results similar to Theorems 5 and 6, giving "proofs" involving generalized Riccati transformations. The proof of his Corollary 2.2, for example requires $U \neq 0$ throughout $\bar{G}$ ($v \neq 0$ in our notation) in order that his Corollary 2.1 can be applied; thus Corollary 2.2 is not correctly proved and the proof cannot be repaired without additional machinery (like Lemma 4 above). In the one-dimensional case, of course, as described in Theorem 2, the boundary limits can be found from L'Hôpital's rule since solutions v of (6) are known to have only simple zeros (by the uniqueness of the solution of an initial value problem). In higher dimensions, two alternative procedures have been indicated in the proofs of Theorems 5 and 6. Wong also assumes throughout that ∂G is piecewise smooth: This is too much when u is identically zero on ∂G, as in Theorem 5, and too little otherwise, as in Theorem 6. We remark that the case of F in (24) being finite on a proper subboundary Γ_2 of ∂G and $F=\infty$ on the complementary subboundary Γ_1 (i.e. $v \equiv 0$ on Γ_1) remains unsettled: Wong's proof is not valid.

The case of unbounded domains G in E^n can be handled by techniques similar to those used by the author in [7]; one has boundary limits analogous to the limits (3) replacing the boundary integral in (25), yielding an n-dimensional version of the inequality (4).

REFERENCES

[1] P. R. BEESACK, *Integral inequalities of the Wirtinger type*, Duke Math. J., 25 (1958), pp. 477–498.

[2] J. B. DIAZ AND F. T. METCALF, *Variations of Wirtinger's inequality*, Inequalities, Academic Press, New York, 1967, pp. 79–103.

[3] K. FAN, O. TAUSSKY AND J. TODD, *Discrete analogs of inequalities of Wirtinger*, Monatsh. Math., 59 (1955), pp. 73–90.

[4] G. H. HARDY, J. E. LITTLEWOOD AND G. PÓLYA, *Inequalities*, 2nd ed., Cambridge University Press, London and New York, 1959.

[5] J. G. HEYWOOD, E. S. NOUSSAIR AND C. A. SWANSON, *On the zeros of solutions of elliptic inequalities in bounded domains*. J. Differential Equations, to appear.

[6] C. A. SWANSON, *A comparison theorem for elliptic differential equations*, Proc. Amer. Math. Soc., 17 (1966), pp. 611–616.

[7] ———, *An identity for elliptic equations with applications*, Trans. Amer. Math. Soc., 134 (1968), pp. 325–333.

[8] ———, *Comparison theorems for elliptic systems*, Pacific J. Math., 33 (1970), 445–450.

[9] P.-K. WONG, *Wirtinger type inequalities and elliptic differential inequalities*, Tôhoku Math. J., 23 (1971), pp. 117–127.

SIAM J. MATH. ANAL.
Vol. 9, No. 3, June 1978

Copyright © 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0903-0005$1.00/0

AN IMPLICIT FUNCTION THEOREM AND ITS APPLICATIONS TO NONLINEAR ELECTRICAL NETWORKS*

T. MATSUMOTO†

Abstract. An implicit function theorem is given for the equation $G(x, y, t)=0$, where G is continuously differentiable in (x, y) and measurable in t. Also discussed is the differential–algebraic system $dx/dt = F(y, t)$, $G(x, y, t)=0$, where G is defined above and F is continuously differentiable in y and measurable in t. It is shown that the results are useful in the analysis of nonlinear electrical networks.

1. Introduction. During the course of a study of electrical networks, the author was naturally led to differential-algebraic systems of the form

$$\frac{dx}{dt}=F(y, t), \tag{1}$$

$$G(x, y, t)=0 \tag{2}$$

where F is an $\mathbb{R}^n$-valued function defined on $\mathbb{R}^m \times \mathbb{R}$ and G is an $\mathbb{R}^m$-valued function defined on $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$. In order that the differential-algebraic equation (1)–(2) be well defined, there must first exist a y satisfying (2) for a given (x, t). Such a y, in circuit theory, is called an *operating point* at (x, t). Secondly, a single-valued function $y(x, t)$ must be determined via (2). In electrical networks G is often discontinuous in t so that the problem is nontrivial. We will first give conditions under which there is a unique operating point to (2). Then, supposing that G and G_x, G_y are continuous in (x, y) and measurable in t, we will state and prove an implicit function theorem for (2). We will also indicate how the results can be applied to electrical networks.

2. Results. Let G be an $\mathbb{R}^m$-valued function defined on $\mathbb{R}^n \times \mathbb{R}^m \times I$, where I is a compact interval.

Assumption 1. $G(\cdot, \cdot, t)$ and $G_x(\cdot, \cdot, t)$, $G_y(\cdot, \cdot, t)$ are continuous and $G(x, y, \cdot)$ and $G_x(x, y, \cdot)$, $G_y(x, y, \cdot)$ are (Borel) measurable.

We will first give conditions under which for a given (x, t), there is a unique $y(x, t)$ satisfying

$$G(x, y(x, t), t)=0. \tag{2.1}$$

PROPOSITION 1. *Let (x, t) be given and suppose that the following hold:*

(A) $$\lim_{\|y\|\to\infty} (\|y\|\,\|G(x, y, t)\|+\langle y, G(x, y, t)\rangle)>0.$$

(B) *For each y*

$$\det [G_y(x, y, t)]\neq 0.$$

Then there is a unique $y(x, t)$ satisfying (2.1).

Proof. It follows from condition (A) that there is a bounded open subset D of $\mathbb{R}^m$ such that $G(x, \cdot, t)$ has no zeros in $\mathbb{R}^m - D$ and for every y in the boundary ∂D of D,

$$\|y\|\,\|G(x, y, t)\|+\langle y, G(x, y, t)\rangle>0.$$

* Received by the editors February 19, 1976, and in final revised form November 12, 1976.

† Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, California, 94720. Presently on leave from Department of Electrical Engineering, Waseda University, Tokyo 160, Japan. This work was supported by the Kawakami Memorial Foundation.

Geometrically, this means that for every y in ∂D, the vectors y and $G(x, y, t)$ never point in opposite directions. Hence the homotopy

$$\mathcal{H}(y, \tau) = \tau y + (1-\tau)G(x, y, t) \tag{2.2}$$

has no zeros on $\partial D \times [0, 1]$. In order to prove the statement we will make use of some properties of the degree of a function. If a continuously differentiable function f from $\mathbb{R}^m$ into itself has zeros in a domain D of $\mathbb{R}^m$, the degree is defined by

$$\text{degree}\,(f, 0, D) = \sum_{y \in A} \text{sgn}\det[D_y f(y)]$$

where $A = \{y \in \mathbb{R}^m | f(y) = 0\}$. The determinant is assumed nonzero at each point of A. If for some y in A, $\det[D_y f(y)] = 0$, then pick a sequence $\{p_k\}$ in $\mathbb{R}^m$ with $p_k \to 0$ such that for all y in $A_k = \{y \in \mathbb{R}^m | f(y) = p_k\}$, $\det[D_y f(y)] \neq 0$, $k = 1, 2, \cdots$. Such a sequence exists because the set $\{y \in \mathbb{R}^m | \det[D_y f(y)] = 0\}$ has Lebesgue measure zero. Finally one defines

$$\text{degree}\,(f, 0, D) = \lim_{k \to \infty} \text{degree}\,(f, p_k, D)$$

where

$$\text{degree}\,(f, p_k, D) = \sum_{y \in A_k} \text{sgn}\det[D_y f(y)].$$

It can be shown that this limit process is well defined [1]. One of the important properties of the degree is the homotopy invariance. Namely, if $\mathcal{H}(y, \tau)$ defined by (2.2) has no zeros on $\partial D \times [0, 1]$, then

$$\text{degree}\,(i_d, 0, D) = \text{degree}\,(G(x, \cdot, t), 0, D) \tag{2.3}$$

where i_d is the identity map of $\mathbb{R}^m$. Clearly, there is no loss of generality in assuming that D contains the origin of $\mathbb{R}^m$ so that

$$\text{degree}\,(i_d, 0, D) = 1. \tag{2.4}$$

It follows from this and (2.3) that

$$\text{degree}\,(G(x, \cdot, t), 0, D) = 1.$$

If a function has a nonzero degree, it has at least one zero [1] so that there is at least one y satisfying (2.1).

In order to prove uniqueness set

$$\mathcal{Y}(x, t) = \{y \in \mathbb{R}^m | G(x, y, t) = 0\}.$$

It follows from condition (B) that $\det[G_y(x, y, t)]$ never changes sign for all y. Hence

$$\begin{aligned} 1 = \text{degree}\,(G(x, \cdot, t), 0, D) &= \sum_{y \in \mathcal{Y}(x, t)} \text{sgn}\det[G_y(x, y, t)] \\ &= \text{number of elements in } \mathcal{Y}(x, t), \end{aligned} \tag{2.5}$$

which proves uniqueness.

Remarks. Note that conditions (A) and (B) imply $\text{sgn}\det[G_y(x, y, t)] = +1$. If condition (A) is replaced by

$$\lim_{\|y\| \to \infty} (\|y\|\,\|G(x, y, t)\| - \langle y, G(x, y, t)\rangle) > 0 \tag{A*}$$

then

$$\text{degree}\,(G(x,\cdot,t),0,D)=\text{degree}\,(-i_d,0,D)=-1$$

so that sgn det $[G_y(x, y, t)] = -1$ holds. If conditions (A*) and (B) hold, then

$$\begin{aligned}-1=\text{degree}\,(G(x,\cdot,t),0,D)&=\sum_{y\in\mathcal{Y}(x,t)}\text{sgn}\det[G_y(x,y,t)]\\&=(-1)\times\text{number of elements in }\mathcal{Y}(x,t).\end{aligned}\tag{2.6}$$

Hence the existence and uniqueness is guaranteed by (A*) and (B) also.

The following is a sufficient condition for (A).

$$\lim_{\|y\|\to\infty}\langle y, G(x,y,t)\rangle>0.\tag{A$'$}$$

This condition can be considered as a generalization of the concept of *passivity* in electrical networks. Namely, $G(x,\cdot,t)$ is passive at (x,t) if $\langle y, G(x,y,t)\rangle>0$ for all y. Condition (A$'$) might be called *eventual passivity* of $G(x,\cdot,t)$ at (x,t) in the sense that $\langle y, G(x,y,t)\rangle>0$ outside a bounded subset of $\mathbb{R}^m$.

A result related to this problem is obtained in [2], where it is shown that if

$$\lim_{\|y\|\to\infty}\|G(x,y,t)\|=\infty\tag{2.7}$$

and if condition (B) of Proposition 1 holds, then there is a unique $y(x,t)$ satisfying (2.1). There are functions that satisfy (A) and (B) while they do not satisfy (2.7) and (B). Consider, for example,

$$G(x,y,t)=x+\text{Arctan}\,y+t$$

where all variables are scalar and Arctan y is the principal value of arctan y. Clearly, this function does not satisfy (2.7). However, for $-\pi/2<x+t<\pi/2$, $G(x,\cdot,t)$ satisfies condition (A). Furthermore

$$G_y(x,y,t)=1/[1+(y)^2]\neq 0$$

so that condition (B) is satisfied. It should be noted that (2.7) does not necessarily imply (A$'$), so that condition (2.7) alone does not imply the existence of at least one zero of (2). For example

$$G(x,y,t)=x+(y)^2+t$$

satisfies (2.7) for every (x,t) but it does not satisfy (A$'$) for any (x,t). Clearly this function may not have zeros.

We will next give our implicit function theorem.

PROPOSITION 2. *If for every (x,t) conditions (A) and (B) of Proposition 1 are satisfied, then there is a unique function $y(\cdot,\cdot)$ on $\mathbb{R}^n\times I$ into $\mathbb{R}^m$ satisfying*

$$G(x,y(x,t),t)=0.$$

Furthermore

(i) *$y(\cdot,t)$ is continuous and $y(x,\cdot)$ is (Lebesgue) measurable,*

(ii) *$y_x(x,t)$ exists and is given by*

$$y_x(x,t)=-[G_y(x,y(x,t),t)]^{-1}G_x(x,y(x,t),t).$$

A fortiori the left hand side is continuous in x and (Lebesgue) measurable in t.

For the proof, the Scorza Dragoni lemma [3] turns out to be useful.

LEMMA (Scorza Dragoni). *Fix x. For every $\varepsilon > 0$ there is a closed subset I_ε of I with the following properties.*

(α) *$G(x, \cdot, \cdot)$ is jointly continuous on $\mathbb{R}^m \times I_\varepsilon$.*

(β) *$\mu(I - I_\varepsilon) < \varepsilon$, where μ is the Lebesgue measure on $\mathbb{R}$.*

Proof of Proposition 2. By Proposition 1, $y(\cdot, \cdot)$ is well defined. In order to prove (i) of Proposition 2 we first fix x and show the measurability of $y(x, \cdot)$. It follows from the lemma that for an arbitrary positive integer i, there is a closed subset I_i of I such that $G(x, \cdot, \cdot)$ is jointly continuous on $\mathbb{R}^m \times I_i$ and $\mu(I - I_i) < 1/i$. Define $I_{ik} = \{t \in I_i \mid \|y(x, t)\| \leqq k\}$. Clearly, then $I_i = \bigcup_{k=1}^{\infty} I_{ik}$. We claim that $y(x, \cdot)$ is continuous on I_{ik}. To show this let $t_n \to t$, $t_n, t \in I_{ik}$. Then the corresponding sequence $\{y(x, t_n)\}$ satisfies

$$\|y(x, t_n)\| \leqq k, \qquad n = 1, 2, \cdots,$$

so that there is a subsequence $\{y(x, t_{n_j})\}$ with

$$y(x, t_{n_j}) \to y^* \in \mathbb{R}^m.$$

It follows from the definition of I_i that $G(x, \cdot, \cdot)$ is jointly continuous on $\mathbb{R}^m \times I_{ik}$ so that

$$\lim_{j \to \infty} G(x, y(x, t_{n_j}), t_{n_j}) = G(x, y^*, t) = 0.$$

Since $y(x, t)$ is unique, we must have $y^* = y(x, t)$. Hence $y(x, \cdot)$ is continuous on I_{ik} so that it is measurable on the countable union

$$\bigcup_{i=1}^{\infty} \bigcup_{k=1}^{\infty} I_{ik} = \bigcup_{i=1}^{\infty} I_i.$$

Finally noticing that $\bigcup_{i=1}^{n} I_i$ is increasing in n, we see that

$$\mu\left(I - \bigcup_{i=1}^{\infty} I_i\right) = \lim_{n \to \infty} \mu\left(I - \bigcup_{i=1}^{n} I_i\right) \leqq \lim_{n \to \infty} \frac{1}{n} = 0.$$

Hence $y(x, t)$ is (Lebesgue) measurable in t on I.

The property with respect to x and (ii) follow from a standard implicit function theorem.

It should be noted that in [2], $G(x, y, \cdot)$ is assumed continuous.

Given the fact that $y(x, t)$ is determined by (2), one can substitute it into (1):

$$dx/dt = F(y(x, t), t) \tag{2.8}$$

so that various techniques in differential equations can be used.

3. Applications. In this section we will show how the results of the previous section can be applied to electrical networks. An electrical network is an interconnection, however complicated, of elements of four basic kinds. They are resistors, capacitors, inductors and independent sources. Let ρ, γ and λ be the number of resistors, capacitors and inductors, respectively. Except for independent sources, there are four kinds of variables in a network. They are capacitor charges (denoted by q_C), inductor fluxes (φ_L), branch voltages (v) and branch currents (i). Given a network, a unique linear graph is defined which shows the topology of the network. A tree for the linear graph is called a *proper tree* [4] if it contains all the capacitors and all the independent voltage sources in the network and the cotree contains all the inductors and all the independent current sources in the network. For the sake of simplicity we assume that a proper tree exists. It is easy to relax this assumption. Pick a proper tree and partition branch voltages and branch currents as

$$v=(v_R, v_L, v_C, v_G) \quad \text{and} \quad i=(i_R, i_L, i_C, i_G)$$

respectively, where R, L, C and G denote link resistors, inductors, capacitors and tree branch resistors, respectively. There are three constraints that must be satisfied by a network.

Constraint 1. *Kirchhoff laws*:

$$v_R=-F_{RC}v_C-F_{RC}v_G+e_R(t), \tag{3.1}$$

$$v_L=-F_{LC}v_C-F_{LG}v_G+e_L(t), \tag{3.2}$$

$$i_C=F_{RC}^T i_R+F_{LC}^T i_L+j_C(t), \tag{3.3}$$

$$i_G=F_{RG}^T i_R+F_{LG}^T i_L+j_G(t), \tag{3.4}$$

where F_{RC}, F_{RG} etc. are matrices that are uniquely determined once a tree is chosen. $e_R(t)$ and $e_L(t)$ are independent voltage sources and $j_C(t)$ and $j_G(t)$ are independent current sources. T denotes the matrix transposition.

Constraint 2. *Branch characteristics*:

$$f(v_R, v_G, i_R, i_G, v_C, i_C, v_L, i_L, \nu, t)=0 \tag{3.5}$$

$$g(v_C, q_C, \mu, t)=0 \tag{3.6}$$

$$h(i_L, \varphi_L, \eta, t)=0 \tag{3.7}$$

where f, g and h take values in $\mathbb{R}^\rho$, $\mathbb{R}^\gamma$ and $\mathbb{R}^\lambda$, respectively, ν, μ and η are parameters (capacitance, inductance, temperature etc.), and t is time.

Note that dependent sources [4] are included in (3.5) and that couplings among elements of different kinds are allowed so that (3.5)–(3.7) cover a very general class of network including transistors, vacuum tubes and various electronic devices.

Constraint 3. *Maxwell's equations*:

$$dq_C/dt=i_C, \qquad d\varphi_L/dt=v_L. \tag{3.8}$$

In order to put these constraints into the form described in § 2, substitute (3.1)–(3.4) into (3.5) and obtain

$$f^*(v_G, i_R, v_C, i_L, \nu, t)=0. \tag{3.9}$$

Next set $x=(q_C, \varphi_L)$, $y=(v_G, i_R, v_C, i_L)$ and $\xi=(\nu, \mu, \eta)$. Then (3.9), (3.6) and (3.7) can be written as

$$G(x, y, \xi, t)=0. \tag{3.10}$$

Note that y is a $\rho+\gamma+\lambda$ vector and that G takes its values in $\mathbb{R}^{\rho+\gamma+\lambda}$. It follows from (3.2), (3.3) and Maxwell's equations (3.8) that

$$\frac{dx}{dt}=\begin{bmatrix} i_C \\ v_L \end{bmatrix}=Wy+\begin{bmatrix} j_C(t) \\ e_L(t) \end{bmatrix}\equiv F(y, t), \tag{3.11}$$

where

$$W=\begin{bmatrix} 0 & F_{RC}^T & 0 & F_{LC}^T \\ -F_{LG} & 0 & -F_{LC} & 0 \end{bmatrix}. \tag{3.12}$$

Thus (3.10)–(3.11) is the differential-algebraic system discussed in § 2.

The functions F and G of (3.10)–(3.11) are sometimes discontinuous in t. This happens, for example, when square pulses are applied to a network and when the branch characteristics are discontinuous in t. The first question in electrical network is the existence and uniqueness of $y(x, t)$ satisfying (2.1), i.e., the operating point. The next question is the existence and uniqueness of solutions to (3.10)–(3.11) given the fact that $y(\cdot, \cdot)$ is well defined. Another important problem is the calculation of the derivative $\partial x/\partial \xi$ which is called the *sensitivity of x with respect to ξ*. We will show that results of § 2 are useful for the above problems.

PROPOSITION 3. *Consider the network described above:*

$$\frac{dx}{dt} = Wy + \begin{bmatrix} j_C(t) \\ e_L(t) \end{bmatrix} \equiv F(y, t) \tag{3.11}$$

$$G(x, y, \xi, t) = 0. \tag{3.10}$$

Let Assumption 1 *be satisfied with* (x, y) *replaced by* (x, y, ξ) *and let* $j_C(t)$ *and* $e_L(t)$ *be measurable. Suppose that for every* (x, ξ, t), the function $G(x, \cdot, \xi, t)$ *satisfies conditions* (A) *and* (B) *of Proposition* 1.

Let

$$B(x, \xi, t) = \begin{bmatrix} 0 & 0 \\ \partial g/\partial q_C & 0 \\ 0 & \partial h/\partial \varphi_L \end{bmatrix}.$$

If for a given (x_0, ξ_0, t_0) *there is a region* $X \times \Lambda \times J$ *containing* (x_0, ξ_0, t_0) *and there is an integrable function* $m(\cdot)$ *on* J *with*

$$\|F(y(x, \xi, t), t)\| \leqq m(t), \qquad \|[G_y(x, y(x, \xi, t), \xi, t)]^{-1}B(x, \xi, t)\| \leqq m(t)$$
$$(x, \xi, t) \in X \times \Lambda \times J,$$

then there is a unique solution to (3.10)–(3.11) *on a nonvanishing interval containing* t_0, *for each* ξ. *Furthermore, the solution is continuous in* ξ.

Proof. It follows from Proposition 1 that $y(\cdot, \cdot, t)$ and $y_x(\cdot, \cdot, t)$ are continuous and $y(x, \xi, \cdot)$ and $y_x(x, \xi, \cdot)$ are measurable. Note that by (3.11)

$$F_x(y(x, \xi, t), t) = Wy_x(x, \xi, t).$$

It follows from (ii) of Proposition 2 that

$$y_x(x, \xi, t) = -[G_y(x, y(x, \xi, t), \xi, t)]^{-1}B(x, \xi, t). \tag{3.13}$$

Hence if the above inequalities are satisfied, then $\|F\|$ and $\|F_x\|$ are dominated by an integrable function. The result follows from a standard existence theorem for differential equations. See [5], for example.

PROPOSITION 4. *Let* $D(x, \xi, t)$ *be the block diagonal matrix* $\operatorname{diag}(f_\nu^*, g_\mu, h_\eta)$. *Suppose that the conditions of Proposition* 3 *are satisfied. If, in addition, the norm*

$$\|[G_y(x, y(x, \xi, t), \xi, t)]^{-1}D(x, \xi, t)\|, \qquad (x, \xi, t) \in X \times \Lambda \times J$$

is dominated by an integrable function on J, *then* x *is differentiable with respect to* ξ. *Furthermore* $\partial x/\partial \xi \equiv \psi$ *(the sensitivity of* x *with respect to* ξ*) satisfies the following linear differential equation for almost every* t:

$$d\psi/dt = -W[G_y]^{-1}[B\psi + D] \tag{3.14}$$

where B *is defined in the statement of Proposition* 3 *and* W *is defined by* (3.12).

Proof. It follows from (3.11) that

$$F_\xi(y(x,\xi,t),t)=Wy_\xi(x,\xi,t). \tag{3.15}$$

Again, by (ii) of Proposition 2,

$$y_\xi(x,\xi,t)=-[G_y(x,y(x,\xi,t),\xi,t)]^{-1}D(x,\xi,t). \tag{3.16}$$

Now, if the conditions of Proposition 4 are satisfied, then (3.15) and (3.16) imply that $\|F_\xi\|$ is dominated by an integrable function. Hence a result in differential equation (see [5], for example) tells us that x is differentiable with respect to ξ and $\partial x/\partial\xi\equiv\psi$ satisfies

$$y_\xi(x,\xi,t)=-[G_y(x,y(x,\xi,t),\xi,t)]^{-1}D(x,\xi,t). \tag{3.16}$$

Formula (3.14) follows from substitution of (3.13) and (3.16) into (3.17).

Remarks. There are many nontrivial networks that satisfy the conditions of Proposition 1. *Uniformly positive definite networks* [6], for example, satisfy those conditions. The differential equation (3.14) is called the *sensitivity equation.* The sensitivity $\partial x/\partial\xi$ plays many important roles in electrical networks. (3.14) is a fairly explicit formula which has not yet been obtained. When one solves (3.14) numerically, one of the most time consuming operations would be the one involving $[G_y]^{-1}$. It should be observed, however, that an algorithm involving this matrix is already available. Because in the process of solving (3.10), one often uses the Newton–Raphson iteration:

$$[y]_{k+1}=[y]_k-[G_y]_k^{-1}[G]_k$$

where $[G_y]_k^{-1}$ and $[G]_k$ mean that $[G_y]^{-1}$ and G are evaluated at $[y]_k$. It should also be noted that G_y has a relatively simple form:

$$G_y=\left[\begin{array}{cc|c|c} \partial f^*/\partial v_G & \partial f^*/\partial i_R & \partial f^*/\partial v_C & \partial f^*/\partial i_L \\ \hline \multicolumn{2}{c|}{0} & \partial g/\partial v_C & 0 \\ \hline \multicolumn{2}{c|}{0} & 0 & \partial h/\partial i_L \end{array}\right].$$

Note that for each (ξ,t), the set $\Sigma(\xi,t)=\{(x,y)|G(x,y,\xi,t)=0\}$ can be viewed as a manifold under certain assumptions. Hence (3.10) and (3.11) will define a vector field on a manifold. In [7], [8] and [9] electrical networks are looked at from that point of view.

Another successful application of the degree theory to electrical networks is found in [10], where operating points of resistive networks are discussed.

Acknowledgment. The author is indebted to Prof. Y. Ishizuka and the reviewers for comments.

REFERENCES

[1] M. BERGER AND M. BERGER, *Perspectives in Nonlinearity*, W. A. Benjamin, New York, 1968.
[2] M. IKEDA AND S. KODAMA, *Large-scale dynamical systems*, IEEE Trans. Circuit Theory, CT-20 (1973), pp. 193–202.
[3] M. JACOBS, *Remarks on some recent extensions of Fillipov's implicit functions lemma*, SIAM J. Control, 5 (1967), pp. 622-627.
[4] R. ROHRER, *Circuit Theory*, McGraw-Hill, New York, 1970.
[5] E. LEE AND L. MARKUS, *Foundations of Optimal Control Theory*, John Wiley, New York, 1968.
[6] T. OHTSUKI AND H. WATANABE, *State variable analysis of RLC networks containing nonlinear coupling elements*, IEEE Trans. Circuit Theory, CT-16 (1969), pp. 26–38.

[7] S. SMALE, *On the mathematical foundations of electrical circuit theory*, J. Differential Geometry, 7 (1972), pp. 193–210.

[8] T. MATSUMOTO, *On the dynamics of electrical networks*, J. Differential Equations, 21 (1976), pp. 179–196.

[9] T. MATSUMOTO, *On several geometric aspects of nonlinear networks*, J. Franklin Inst. (Special Issue), 301 (1976), pp. 203–225.

[10] T. OHTSUKI, T. FUJISAWA AND S. KUMAGAI, *Existence theorems and a solution algorithm for piecewise-linear networks*, this Journal, 8 (1977), pp. 69–99.

SIAM J. MATH. ANAL.
Vol. 9, No. 3, June 1978

Copyright © 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0903-0006$01.00/0

A POSITIVE POLYNOMIAL EXPANSION PROBLEM IN MINIMUM VARIANCE SMOOTHING*

LEO W. LAMPONE†

Abstract. In studying stability properties in minimum variance smoothing of discrete data, one is led to consider positive expansions of one set of polynomials in terms of another. Both sets of polynomials are characterized by orthogonality and normalization conditions with respect to different distributions on $(0, \infty)$. We show positivity in the case where the distributions are related by $dv(x) = (x+\alpha)^c\, du(x), c > 0, \alpha \geqq 0$, generalizing a recently proved conjecture of R. Askey. In an application, stability of a large class of minimum variance smoothing formulas is established.

1. Introduction. Consider the problem of smoothing a sequence of observations

$$v_r = f(r) + \varepsilon_r, \qquad r = 0, \pm 1, \pm 2, \cdots, \tag{1.1}$$

where f is a fixed but unknown polynomial of degree less than or equal to $2k$ and $\{\varepsilon_r\}$ is a sample sequence from a real-valued stationary time series with zero mean and continuous spectral density

$$\Phi(\lambda) = \sum_{r=-\infty}^{\infty} \phi_r \cos(r\lambda) \qquad (\phi_{-r} = \phi_r).$$

We apply to (1.1) a symmetric moving average

$$u_r = \sum_{s=-n}^{n} w_s v_{r-s}, \qquad r = 0, \pm 1, \pm 2, \cdots, \tag{1.2}$$

where the coefficient vector

$$W = \begin{bmatrix} w_{-n} \\ \cdot \\ \cdot \\ \cdot \\ w_0 \\ \cdot \\ \cdot \\ \cdot \\ w_n \end{bmatrix} \tag{1.3}$$

satisfies

$$w_s = w_{-s}, \qquad -n \leqq s \leqq n, \tag{1.4a}$$

$$\sum_{s=-n}^{n} w_s s^{2j} = \delta_{0,j}, \qquad 0 \leqq j \leqq k. \tag{1.4b}$$

Thus, the output from (1.2) is $u_r = f(r) + \xi_r$ where

$$\xi_r = \sum_{s=-n}^{n} w_s \varepsilon_{r-s}, \qquad r = 0, \pm 1, \pm 2, \cdots.$$

The process $\{\xi_r\}$ is also stationary with zero mean and variance

$$\sigma^2(W) = \sum_{r,s=-n}^{n} \phi_{r-s} w_r w_s. \tag{1.5}$$

* Received by the editors September 2, 1976, and in revised form November 29, 1976.

† Department of Mathematics, Spring Garden College, Chestnut Hill, Pennsylvania 19118.

If $n>k$ and $\{\varepsilon_r\}$ is of rank greater than or equal to $2n+1$, then (1.5) is a positive-definite form and consequently there is a unique vector $\hat{W}$ satisfying (1.4) for which $\sigma^2(W)$ is minimal. The smoothing formula (1.2) which results from $\hat{W}$ is denoted by Trench in [8] as $MV(n,k;\Phi)$; this stands for "minimum variance smoothing formula, with respect to Φ, of span $2n+1$ and degree $2k+1$."

Considerable research has been devoted to this estimation problem. The trigonometric polynomial

$$C(\theta)=\sum_{s=-n}^{n} w_s \cos(s\theta),$$

is called the characteristic function of the smoothing operation (1.2). Schoenberg [5] has given reasons for calling a smoothing formula (1.2) stable if $C(\theta)$ satisfies

$$|C(\theta)|<1, \qquad 0<\theta\leqq\pi. \tag{1.6}$$

We thus observe that $\sigma^2(W)$ can be written

$$\sigma^2(W)=\frac{1}{2\pi}\int_{-\pi}^{\pi} C^2(\theta)\Phi(\theta)\,d\theta. \tag{1.7}$$

In [8], Trench showed that by making the substitution $x=\sin^2(\theta/2)$, the problem of minimizing (1.7) subject to (1.4) can be reformulated as minimizing the integral

$$\frac{1}{\pi}\int_0^1 p^2(x)F(x)\,dx \tag{1.8}$$

over the convex set of polynomials of the form

$$p(x)=1-\sum_{j=k+1}^{n} b_j x^j. \tag{1.9}$$

In (1.8), $F(x)=x^{-1/2}(1-x)^{-1/2}\psi(x)$, where $\psi(\sin^2(\theta/2))=\Phi(\theta)$. He showed that the minimizing polynomial $\hat{p}_n(x)$ is the unique polynomial of the form (1.9) satisfying

$$\int_0^1 p(x)x^jF(x)\,dx=0, \qquad k+1\leqq j\leqq n. \tag{1.10}$$

Also, $\hat{p}_n(\sin^2(\theta/2))$ is the characteristic function of the smoothing operation $MV(n,k;\Phi)$. Suppose that $\Gamma(\theta)=[\sin^2(\theta/2)]^{c_1}[1+\lambda\sin^2(\theta/2)]^{c_2}\Phi(\theta)$ where λ, c_1, c_2 are nonnegative numbers. We ask if stability of $MV(n,k;\Phi)$ implies stability of $MV(n,k;\Gamma)$. Note that replacing $\Phi(\theta)$ by $\Gamma(\theta)$ in (1.7) is equivalent to replacing $F(x)$ by $x^{c_1}[1+\lambda x]^{c_2}F(x)$ in (1.8). We consequently write

$$\hat{q}_n(x)=\sum_{j=k}^{n} a_{jn}^{(k)}\hat{p}_j(x)$$

where $\hat{p}_k(x)\equiv 1$ and $\hat{q}_n(\sin^2(\theta/2))$ is the characteristic function of $MV(n,k;\Gamma)$. Thus, if $D_{nk}(\theta)$ represents the characteristic function of $MV(n,k;\Gamma)$ and $C_{jk}(\theta)$ the characteristic functions of $MV(j,k;\Phi)$, $k<j\leqq n$, and $C_{kk}(\theta)\equiv 1$, then

$$D_{nk}(\theta)=\hat{q}_n\left(\sin^2\left(\frac{\theta}{2}\right)\right)=\sum_{j=k}^{n} a_{jn}^{(k)}\hat{p}_j\left(\sin^2\left(\frac{\theta}{2}\right)\right)=\sum_{j=k}^{n} a_{jn}^{(k)}C_{jk}(\theta). \tag{1.11}$$

Since $\sum_{j=k}^{n} a_{jn}^{(k)}=1$, we see that the condition $a_{jn}^{(k)}>0$, $k\leqq j\leqq n$, is sufficient to guarantee stability of $MV(n,k;\Gamma)$ given stability of $MV(j,k;\Phi)$, $j\geqq k+1$. This condition will follow from a general result which we now state as Theorem 1.1. The balance of this paper will then be devoted to the proof of Theorem 1.1, with its application to the smoothing problem being summarized in § 4.

THEOREM 1.1. *Suppose $du(x)$ is a distribution on $(0,\infty)$ and k, is a fixed non-negative integer. Define $p_{kk}(x)=1$. For $n>k$, let $p_{nk}(x)$ be a polynomial of degree n defined by the conditions*:

$$\int_0^\infty p_{nk}(x)x^j\,du(x)=0, \qquad 0\leqq j\leqq n-k-1, \tag{1.12}$$

and

$$p_{nk}^{(r)}(0)=\delta_{0,r}, \qquad 0\leqq r\leqq k. \tag{1.13}$$

Define $q_{nk}(x)$ similarly with respect to the distribution

$$dv(x)=(x+\alpha)^c\,du(x), \qquad c>0, \quad \alpha\geqq 0.$$

Then we can write

$$q_{nk}(x)=\sum_{s=k}^{n} a_{sn}^{(k)}p_{sk}(x) \tag{1.14}$$

where

$$a_{sn}^{(k)}>0, \qquad 0\leqq k\leqq s\leqq n. \tag{1.15}$$

Before beginning the proof of Theorem 1.1 some observations are in order. It is to be understood that $du(x)$ and $dv(x)$ have moments of all orders on $(0,\infty)$ and that $n\leqq N-1$ if $du(x)$ is a discrete distribution over only N points. By definition the polynomials $q_{nk}(x)$ satisfy

$$\int_0^\infty q_{nk}(x)x^j(x+\alpha)^c\,du(x)=0, \qquad 0\leqq j\leqq n-k-1, \tag{1.16}$$

and

$$q_{nk}^{(r)}(0)=\delta_{0,r}, \qquad 0\leqq r\leqq k. \tag{1.17}$$

From (1.13) we see that $p_{nk}(x)$ has the form

$$p_{nk}(x)=1-\sum_{s=k+1}^{n} b_{sn}^{(k)}x^s. \tag{1.18}$$

From (1.12) it is clear that the $b_{sn}^{(k)}$, $k+1\leqq s\leqq n$, are uniquely determined by the system of equations

$$\int_0^\infty x^j\,du(x)=\sum_{s=k+1}^{n} b_{sn}^{(k)}\int_0^\infty x^{s+j}\,du(x), \qquad 0\leqq j\leqq n-k-1.$$

(The coefficient matrix of this system is a Gram matrix.) Moreover, the orthogonality conditions (1.12) imply that $p_{nk}(x)$ has at least $n-k$ zeros in $(0,\infty)$; hence, Descartes' rule of signs implies that the coefficients in (1.18) alternate in sign. It follows that

$$(-1)^{s-k-1}b_{sn}^{(k)}>0, \qquad k+1\leqq s\leqq n. \tag{1.19}$$

The inequality (1.19) in particular shows that $p_{nk}(x)$ is indeed a polynomial of degree n. In a similar fashion, the polynomials $q_{nk}(x)$ are uniquely defined. We point out that Trench obtained (1.19) by a different argument in [8] where he also showed that $p_{nk}(x)$ has exactly $n-k$ zeros (all simple) in $(0,\infty)$.

The sequences $\{p_{n,0}\}$ and $\{q_{n,0}\}$ are orthogonal over $(0,\infty)$ with respect to $du(x)$ and $(x+\alpha)^c\,du(x)$ respectively, and normalized so as to be positive at zero. It was conjectured by Askey [1] and proved by Trench [10] that this implies (1.15) for $k=0$. The sequences $\{p_{n,1}\}$ and $\{q_{n,1}\}$ are specific classes of quasi-orthogonal polynomials. (See [4].)

We now prove a series of lemmas and preliminary theorems with the intention of ultimately displaying the coefficients $a_{sn}^{(k)}$ in (1.14) as the solutions of a suitable linear system. In what follows we assume that $0<c<1$. Once Theorem 1.1 is proved for these values of c, it follows for all positive c by iteration.

2. Properties of the polynomials $p_{nk}(x)$. We begin by establishing a recurrence relation satisfied by the polynomials $p_{nk}(x)$.

THEOREM 2.1. *The polynomials $p_{nk}(x)$ satisfy the recurrence formula*

$$p_{nk}(x)=-B_n^{(k)}p_{n,k-1}(x)+A_n^{(k)}p_{n-1,k-1}(x) \tag{2.1}$$

where $A_n^{(k)}>0$ and $B_n^{(k)}>0$, $1\leqq k<n$. Moreover,

$$(-1)^{n-k}\int_0^\infty x^{n-k}p_{nk}(x)\,du(x)>0, \qquad n\geqq k. \tag{2.2}$$

Proof. Define

$$L_{nk}(x;\alpha)=\alpha p_{n,k-1}(x)+(1-\alpha)p_{n-1,k-1}(x) \tag{2.3}$$

where α is a constant. Then $L_{nk}(x;\alpha)$ is a polynomial of degree n such that $L_{nk}^{(r)}(0;\alpha)=\delta_{0,r}$, $0\leqq r\leqq k-1$. Also,

$$\int_0^\infty L_{nk}(x;\alpha)x^j\,du(x)=0, \qquad 0\leqq j\leqq n-k-1.$$

We now choose α so that $L_{nk}^{(k)}(0;\alpha)=0$; that is, so that

$$\alpha b_{kn}^{(k-1)}+(1-\alpha)b_{k,n-1}^{(k-1)}=0. \tag{2.4}$$

With α chosen in this way, $L_{nk}(x;\alpha)$ satisfies conditions (1.12) and (1.13). Therefore,

$$L_{nk}(x;\alpha)\equiv p_{nk}(x). \tag{2.5}$$

Since (2.5) implies that $\alpha b_{nn}^{(k-1)}=b_{nn}^{(k)}$, we see from (1.19) that $\alpha<0$. This completes the proof of (2.1). Now from (2.1)

$$\begin{aligned}\int_0^\infty x^{n-k}p_{nk}(x)\,du(x)&=A_n^{(k)}\int_0^\infty x^{n-k}p_{n-1,k-1}(x)\,du(x)\\&\quad-B_n^{(k)}\int_0^\infty x^{n-k}p_{n,k-1}(x)\,du(x)\\&=A_n^{(k)}\int_0^\infty x^{n-k}p_{n-1,k-1}(x)\,du(x),\end{aligned} \tag{2.6}$$

by the orthogonality conditions (1.12). Repeated application of (2.6) yields

$$\int_0^\infty x^{n-k}p_{nk}(x)\,du(x)=A_n^{(k)}A_{n-1}^{(k-1)}\cdots A_{n-k+1}^{(1)}\int_0^\infty x^{n-k}p_{n-k,0}(x)\,du(x) \tag{2.7}$$

where $A_{n-k+j}^{(j)}>0$, $1\leqq j\leqq k$. Since

$$\int_0^\infty x^jp_{n-k,0}(x)\,du(x)=0, \qquad 0\leqq j\leqq n-k-1,$$

it follows that

$$\int_0^\infty x^{n-k} p_{n-k,0}(x)\, du(x) = -\frac{1}{b_{n-k,n-k}^{(0)}} \int_0^\infty [p_{n-k,0}(x)]^2\, du(x).$$

Now, (1.19) and (2.7) imply (2.2). This completes the proof of the theorem.

Knowledge of the signs of the coefficients in a certain linear combination of the polynomials $p_{nk}(x)$ will be needed later. This information is contained in the following lemma.

LEMMA 2.1. *Let ν be a nonnegative integer and suppose $0<c<1$. Set $a=\nu+c$ and define $h_{k-1,k}(x;a)=x^\nu(x+\alpha)^c$. For $s \geqq k$ let*

$$h_{sk}(x;a)=h_{s-1,k}(x;a)-\beta_s(a)p_{sk}(x)$$

where

$$\beta_s(a)=\frac{\int_0^\infty h_{s-1,k}(x;a)x^{s-k}\,du(x)}{\int_0^\infty p_{sk}(x)x^{s-k}\,du(x)}. \tag{2.8}$$

Then, for $s \geqq k$

$$\int_0^\infty h_{sk}(x;a)x^j\,du(x)=0, \qquad 0 \leqq j \leqq s-k. \tag{2.9}$$

Also, if $k+1 \leqq \nu < s$ we have $(-1)^{\nu-k-1}\beta_s(a)>0$.

Proof. If we use (1.12), equation (2.9) follows from (2.8) by an easy induction argument. Also, $h_{sk}(x;a)$ has the explicit form

$$h_{sk}(x;a)=x^\nu(x+\alpha)^c-\sum_{j=k}^{s}\beta_j(a)p_{jk}(x).$$

We now assume that $k+1 \leqq \nu < s$. Then the derivative of $h_{sk}(x;a)$ is

$$\begin{aligned} h'_{sk}(x;a) &= x^{\nu-1}(x+\alpha)^{c-1}(ax+\nu\alpha)-x^k\xi_{s-k-1}(x) \\ &= x^k\left[ax^{\nu-k-1}(x+\alpha)^{c-1}\left(x+\frac{\nu\alpha}{c+\nu}\right)-\xi_{s-k-1}(x)\right] \end{aligned}$$

where

$$\xi_{s-k-1}(x)=\sum_{j=k}^{s}\beta_j(a)\frac{p'_{jk}(x)}{x^k}.$$

Equation (2.9) implies that $h_{sk}(x;a)$ has at least $s-k+1$ zeros in $(0,\infty)$. Rolle's theorem implies that $h'_{sk}(x;a)$ has at least $s-k$ zeros in $(0,\infty)$ and that consequently the $(s-k-1)$st derivative of

$$\left[ax^{\nu-k-1}(x+\alpha)^{c-1}\left(x+\frac{\nu\alpha}{c+\nu}\right)-\xi_{s-k-1}(x)\right]$$

has at least one zero in $(0,\infty)$. Let

$$f(x)=ax^{\nu-k-1}(x+\alpha)^{c-1}\left(x+\frac{\nu\alpha}{c+\nu}\right)$$

and $m=\nu-k-1\geqq 0$. Since $(x+\nu\alpha/(c+\nu))=(x+\alpha)-c\alpha/(c+\nu)$, we have

$$f(x)=a\left[\sum_{i=0}^{m}\binom{m}{i}(-\alpha)^{m-i}(x+\alpha)^{c+i-1}\right]\left[(x+\alpha)-\frac{c\alpha}{c+\nu}\right]$$

$$=a\sum_{i=1}^{m}\left[\binom{m}{i-1}+\frac{c}{c+\nu}\binom{m}{i}\right](-\alpha)^{m+1-i}(x+\alpha)^{c+i-1}+c(-\alpha)^{m+1}(x+\alpha)^{c-1}$$
$$+a(x+\alpha)^{c+m}$$

where the summation term is absent if $m=0$. Thus

$$f^{(r)}(x)=(a)r!\sum_{i=1}^{m}\left[\binom{m}{i-1}+\frac{c}{c+\nu}\binom{m}{i}\right](-\alpha)^{m+1-i}\binom{c+i-1}{r}(x+\alpha)^{c+i-1-r}$$
$$+(c)r!(-\alpha)^{m+1}\binom{c-1}{r}(x+\alpha)^{c-1-r}+(a)r!\binom{c+m}{r}(x+\alpha)^{c+m-r}.$$

Now, if $0<c<1$ and $1\leqq i\leqq m<r$ the following inequalities may be verified:

$$(2.10)\quad (-1)^{r-i}\binom{c-1+i}{r}>0,\quad (-1)^{r}\binom{c-1}{r}>0\quad\text{and}\quad (-1)^{r-m-1}\binom{c+m}{r}>0.$$

It then follows that for $x>0$

$$(2.11)\qquad (-1)^{r-m-1}f^{(r)}(x)>0.$$

Let $r=s-k-1$; since $m=\nu-k-1$, our assumptions that $k+1\leqq\nu<s$ and $0<c<1$ imply that the inequalities (2.10) are valid and hence (2.11) yields

$$(2.12)\qquad (-1)^{s-\nu-1}f^{(s-k-1)}(x)>0,\qquad x>0.$$

Our earlier observation that the $(s-k-1)$st derivative of $f(x)-\xi_{s-k-1}(x)$ has at least one zero in $(0,\infty)$ implies that

$$(2.13)\qquad f^{(s-k-1)}(x_0)+\beta_s(a)b_{ss}^{(k)}\cdot s\cdot(s-k-1)!=0$$

for some $x_0\in(0,\infty)$. It follows from (1.19), (2.12) and (2.13) that

$$(-1)^{\nu-k-1}\beta_s(a)>0,\qquad k+1\leqq\nu<s.$$

This completes the proof of Lemma 2.1.

The next lemma generalizes the classical result that a function $f(x)$, continuous on an interval $[a,b]$ and satisfying

$$\int_a^b f(x)x^j\,du(x)=0,\qquad 0\leqq j\leqq n,$$

must have at least $n+1$ zeros in (a,b).

LEMMA 2.2. *Suppose $f(x)$ is continuous in $[0,\infty)$ and*

$$(2.14)\qquad \int_0^\infty f(x)x^j\,du(x)=0,\qquad j=0,k+1,\cdots,n.$$

Then $f(x)$ has at least $n-k+1$ zeros in $(0,\infty)$.

Proof. The proof is by contradiction. Since $\int_0^\infty f(x)\,du(x)=0$, $f(x)$ changes sign at least once in $(0,\infty)$. Suppose $f(x)$ changes sign only at points $x_1,x_2,\cdots,x_m$ where $1\leqq m\leqq n-k$. We first show that there is a polynomial $g(x)$ of the form

$$(2.15)\qquad g(x)=1+\sum_{j=n-m+1}^{n}\alpha_j x^j$$

such that

$$g(x)=(x-x_1)(x-x_2)\cdots(x-x_m)w(x) \tag{2.16}$$

where $w(x)$ is of constant sign in $(0,\infty)$. To this end we impose the conditions $g(x_r)=0$, $1\leqq r\leqq m$, and observe that this gives rise to the system of equations

$$\sum_{j=n-m+1}^{n} \alpha_j(x_r)^j=-1, \qquad 1\leqq r\leqq m. \tag{2.17}$$

This is a square system, the coefficient matrix of which has the form

$$\begin{bmatrix} x_1^{n-m+1} & x_1^{n-m+2} & \cdots & x_1^n \\ x_2^{n-m+1} & x_2^{n-m+2} & \cdots & x_2^n \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ x_m^{n-m+1} & x_m^{n-m+2} & \cdots & x_m^n \end{bmatrix}. \tag{2.18}$$

The determinant of the matrix (2.18) has the form

$$(x_1x_2\cdots x_m)^{n-m+1}\det\begin{bmatrix} 1 & x_1 & \cdots & x_1^{m-1} \\ 1 & x_2 & \cdots & x_2^{m-1} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 1 & x_m & \cdots & x_m^{m-1} \end{bmatrix}, \tag{2.19}$$

which is nonzero since the determinant in (2.19) is a Vandermonde determinant. Thus, the system (2.17) is nonsingular and consequently the coefficients $\{\alpha_j\}$ in (2.15) are uniquely determined. Since Descartes' rule of signs implies that the number of sign changes in the coefficients of $g(x)$ is an upper bound on the positive zeros of $g(x)$, we conclude that $g(x)$ has m zeros $x_1, x_2, \cdots, x_m$ in $(0,\infty)$. Therefore, $g(x)$ has the form (2.16) where $w(x)$ has no roots in $(0,\infty)$. Now,

$$f(x)(x-x_1)(x-x_2)\cdots(x-x_m)w(x)$$

is of constant sign in $(0,\infty)$; but

$$\int_0^\infty f(x)(x-x_1)(x-x_2)\cdots(x-x_m)\,du(x)=0,$$

which is a contradiction. Hence, $f(x)$ must have at least $n-k+1$ roots in $(0,\infty)$, and the proof is complete.

Since the linear span of $\{p_{kk},\cdots,p_{nk}\}$ is the same as that of $\{1,x^{k+1},\cdots,x^n\}$, (2.14) is equivalent to

$$\int_0^\infty f(x)p_{jk}(x)\,du(x)=0, \qquad k\leqq j\leqq n.$$

This is used in the next theorem. This theorem generalizes the device used by Trench [10] to prove the Askey conjecture.

THEOREM 2.2. *If $a>0$ and $a\neq 1,\cdots,a\neq n-k-1$, then*

$$(-1)^{n-k}a(a-1)\cdots(a-n+k+1)\int_0^\infty (x+\alpha)^a p_{nk}(x)\,du(x)>0, \qquad n>k. \tag{2.20}$$

Proof. Define

$$G_{nk}(x;a)=(x+\alpha)^a-\sum_{j=0}^{n-k} d_{jn}(a)x^j,$$

where the coefficients $\{d_{jn}(a)\}$ are determined by the conditions

$$\int_0^\infty G_{nk}(x;a)p_{rk}(x)\,du(x)=0, \qquad k\leqq r\leqq n.$$

This implies that

$$\sum_{j=0}^{n-k} d_{jn}(a)\int_0^\infty x^j p_{rk}(x)\,du(x)=\int_0^\infty (x+\alpha)^a p_{rk}(x)\,du(x), \qquad k\leqq r\leqq n. \tag{2.21}$$

By the orthogonality conditions (1.12), the coefficient matrix of this system is upper triangular. Recalling (2.2) we thus see that the $\{d_{jn}(a)\}$ are uniquely determined. By Lemma 2.2, $G_{nk}(x;a)$ has at least $n-k+1$ roots in $(0,\infty)$. This implies that

$$a(a-1)\cdots(a-n+k+1)(x_1+\alpha)^{a-n+k}-d_{n-k,n}(a)(n-k)!=0$$

for some $x_1\in(0,\infty)$. Therefore, if $a\neq 0,\cdots,a\neq n-k-1$,

$$a(a-1)\cdots(a-n+k+1)d_{n-k,n}(a)>0. \tag{2.22}$$

Setting $r=n$ in (2.21) we see that

$$d_{n-k,n}(a)\int_0^\infty x^{n-k}p_{nk}(x)\,du(x)=\int_0^\infty (x+\alpha)^a p_{nk}(x)\,du(x). \tag{2.23}$$

Now (2.2) and (2.22) complete the proof of the theorem.

3. Proof of Theorem 1.1. Let $\phi_0,\phi_1,\cdots,\phi_n,\cdots$ be a sequence of polynomials orthonormal over $(0,\infty)$ with respect to $du(x)$ and normalized so that $\phi_r(0)>0$ for $r\geqq 0$. (If $du(x)$ has only N points of increase in $(0,\infty)$, then we have only $N-1$ of these orthonormal polynomials.) Recalling (1.14), we may write

$$\int_0^\infty q_{nk}(x)\phi_r(x)(x+\alpha)^c\,du(x)=\sum_{j=k}^{n} a_{jn}^{(k)}\int_0^\infty p_{jk}(x)\phi_r(x)(x+\alpha)^c\,du(x), \qquad 0\leqq r\leqq n-k. \tag{3.1}$$

By the orthogonality conditions (1.16) we have

$$\int_0^\infty q_{nk}(x)\phi_r(x)(x+\alpha)^c\,du(x)=\begin{cases}0, & 0\leqq r\leqq n-k-1,\\ \rho_n, & r=n-k.\end{cases}$$

To see that $\rho_n>0$, we note that since $\phi_r(0)>0$, Descartes' rule of signs implies that $(-1)^j\phi_r^{(j)}(0)>0$, $0\leqq j\leqq r$. (As all zeros of $\phi_r(x)$ are in $(0,\infty)$, we actually have $(-1)^j\phi_r^{(j)}(-\alpha)>0$, $\alpha\geqq 0$.) Thus, using (1.16) we can write

$$\begin{aligned}\int_0^\infty q_{nk}(x)\phi_{n-k}(x)(x+\alpha)^c\,du(x)&=\frac{\phi_{n-k}^{(n-k)}(0)}{(n-k)!}\int_0^\infty x^{n-k}q_{nk}(x)(x+\alpha)^c\,du(x)\\ &=\frac{(-1)^{n-k}\phi_{n-k}^{(n-k)}(0)}{(n-k)!}(-1)^{n-k}\int_0^\infty x^{n-k}q_{nk}(x)(x+\alpha)^c\,du(x)>0,\end{aligned} \tag{3.2}$$

from Theorem 2.1. Thus, (3.1) becomes

$$
\begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ \rho_n \end{bmatrix} = \Phi_{nk}(c) \begin{bmatrix} a_{kn}^{(k)} \\ a_{k+1,n}^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ \\ a_{nn}^{(k)} \end{bmatrix} \tag{3.3}
$$

where

$$
\Phi_{nk}(c) = \left[\int_0^\infty \phi_{i-k}(x) p_{jk}(x)(x+\alpha)^c \, du(x) \right]_{k \leqq i,j \leqq n}. \tag{3.4}
$$

We now proceed to record information relative to the entries of $\Phi_{nk}(c)$.

LEMMA 3.1. *For* $0 < c < 1$, *the elements above the diagonal of* $\Phi_{nk}(c)$ *are negative.*

Proof. Setting $a = r + c$ in Theorem 2.2 yields

$$
(-1)^{n-k}(r+c)(r+c-1)\cdots(r+c-n+k+1)\int_0^\infty p_{nk}(x)(x+\alpha)^{r+c}\, du(x) > 0.
$$

If $0 < c < 1$ and $0 \leqq r \leqq n-k-1$ we obtain

$$
(-1)^r \int_0^\infty (x+\alpha)^{r+c} p_{nk}(x)\, du(x) < 0, \qquad 0 \leqq r \leqq n-k-1.
$$

Since

$$
\phi_r(x) = \sum_{j=0}^{r} \frac{\phi_r^{(j)}(-\alpha)}{j!}(x+\alpha)^j
$$

with $(-1)^j \phi_r^{(j)}(-\alpha) > 0$, we see that the elements above the diagonal of $\Phi_{nk}(c)$ satisfy the inequality

$$
\begin{aligned}
&\int_0^\infty p_{sk}(x)\phi_r(x)(x+\alpha)^c\, du(x) \\
&\quad = \sum_{j=0}^{r} \frac{|\phi_r^{(j)}(-\alpha)|}{j!}(-1)^j \int_0^\infty (x+\alpha)^{c+j} p_{sk}(x)\, du(x) < 0, \qquad 0 \leqq r \leqq s-k-1.
\end{aligned}
$$

This completes the proof of the lemma.

We now consider the matrix $\Phi_{nk}(0)$.

LEMMA 3.2. $\Phi_{nk}(0)$ *is a lower triangular matrix with positive elements on the main diagonal and zero elements in the first column below the main diagonal. Moreover, recall* (2.1):

$$
p_{nk}(x) = -B_n^{(k)} p_{n,k-1}(x) + A_n^{(k)} p_{n-1,k-1}(x), \qquad n \geqq k+1,
$$

where $A_n^{(k)} > 0$ *and* $B_n^{(k)} > 0$. *Define* $A_j^{(j)} = 1$ *and* $B_j^{(j)} = 0, j \geqq 1$. *Then, for* $n \geqq k+1$, *the matrices* $\Phi_{nk}(0)$ *satisfy the recurrence formula*

$$\Phi_{nk}(0) = \Phi_{n-1,k-1}(0) \begin{bmatrix} A_k^{(k)} & 0 & 0 & \cdots & 0 & 0 \\ 0 & A_{k+1}^{(k)} & 0 & \cdots & 0 & 0 \\ 0 & -B_{k+1}^{(k)} & A_{k+2}^{(k)} & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & -B_{n-1}^{(k)} & A_n^{(k)} \end{bmatrix}. \tag{3.5}$$

(*The only nonzero elements in the matrix displayed in* (3.5) *occur on the main diagonal and one position below the main diagonal.*)

Proof. The elements above the main diagonal vanish because of the orthogonality conditions (1.12). Elements in the first column below the main diagonal vanish because

$$\int_0^\infty p_{kk}(x)\phi_r(x)\,du(x) = \int_0^\infty \phi_r(x)\,du(x) = 0, \qquad r > 0.$$

As for the elements on the main diagonal,

$$\int_0^\infty \phi_{i-k}(x)p_{ik}(x)\,du(x) > 0, \qquad k \leqq i \leqq n,$$

by a computation similar to (3.2). Now, recall (3.4) (with $c = 0$):

$$\Phi_{nk}(0) = \left[\int_0^\infty \phi_{i-k}(x)p_{jk}(x)\,du(x)\right]_{k \leqq i,j \leqq n}.$$

Thus,

$$\Phi_{n-1,k-1}(0) = \left[\int_0^\infty \phi_{i-k+1}(x)p_{j,k-1}(x)\,du(x)\right]_{k-1 \leqq i,j \leqq n-1}.$$

If we now define

$$\langle p, q\rangle = \int_0^\infty p(x)q(x)\,du(x)$$

the right-hand side of (3.5) is seen to be of the form

$$\begin{bmatrix} \langle \phi_0, p_{k-1,k-1}\rangle & 0 & \cdots & 0 \\ 0 & \langle \phi_1, p_{k,k-1}\rangle & \cdots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & \langle \phi_{n-k}, p_{k,k-1}\rangle & \cdots & \langle \phi_{n-k}, p_{n-1,k-1}\rangle \end{bmatrix}$$

$$\times \begin{bmatrix} A_k^{(k)} & 0 & 0 & \cdots & 0 & 0 \\ 0 & A_{k+1}^{(k)} & 0 & \cdots & 0 & 0 \\ 0 & -B_{k+1}^{(k)} & A_{k+2}^{(k)} & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & -B_{n-1}^{(k)} & A_n^{(k)} \end{bmatrix}. \tag{3.6}$$

The diagonal elements of this product have the form

$$A_i^{(k)}\int_0^\infty \phi_{i-k}(x)p_{i-1,k-1}(x)\,du(x), \qquad k\leqq i, j\leqq n.$$

But

$$\begin{aligned} A_i^{(k)}&\int_0^\infty \phi_{i-k}(x)p_{i-1,k-1}(x)\,du(x) \\ &=\int_0^\infty [A_i^{(k)}p_{i-1,k-1}(x)-B_i^{(k)}p_{i,k-1}(x)]\phi_{i-k}(x)\,du(x) \\ &=\int_0^\infty \phi_{i-k}(x)p_{ik}(x)\,du(x), \qquad k\leqq i\leqq n, \end{aligned}$$

by (2.1) and the orthogonality conditions (1.12).

Now, in the matrix product (3.6) it is clear that elements above the main diagonal and in the first column below the main diagonal are zero. Let a_{ij}, $k<j<i\leqq n$, represent the remaining elements in the product. Then

$$\begin{aligned} a_{ij}&=A_j^{(k)}\int_0^\infty \phi_{i-k}(x)p_{j-1,k-1}(x)\,du(x)-B_j^{(k)}\int_0^\infty \phi_{i-k}(x)p_{j,k-1}(x)\,du(x) \\ &=\int_0^\infty [A_j^{(k)}p_{j-1,k-1}(x)-B_j^{(k)}p_{j,k-1}(x)]\phi_{i-k}(x)\,du(x) \\ &=\int_0^\infty \phi_{i-k}(x)p_{jk}(x)\,du(x), \end{aligned}$$

again with the use of (2.1). Thus, (3.5) is established and the lemma is proved.

If we now set

$$\Gamma_j^{(i)}=\begin{bmatrix} A_i^{(i)} & 0 & 0 & \cdots & 0 & 0 \\ 0 & A_{i+1}^{(i)} & 0 & \cdots & 0 & 0 \\ 0 & -B_{i+1}^{(i)} & A_{i+2}^{(i)} & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & -B_{j-1}^{(i)} & A_j^{(i)} \end{bmatrix} \qquad 1\leqq i\leqq k, \quad j>i,$$

repeated application of (3.5) yields

$$\Phi_{nk}(0)=\Phi_{n-k,0}(0)\Gamma_{n-k+1}^{(1)}\cdots\Gamma_n^{(k)}. \tag{3.7}$$

Since (1.12) implies that $p_{j,0}(x)=\gamma_j\phi_j(x)$ ($\gamma_j>0$), we have

$$\Phi_{n-k,0}(0)=[\gamma_j\delta_{ij}]_{0\leqq i,j\leqq n-k}.$$

Note also that

$$\Gamma_j^{(i)}=\left[\begin{array}{ccc|c} & & & 0 \\ & \Gamma_{j-1}^{(i)} & & \cdot \\ & & & \cdot \\ & & & \cdot \\ & & & 0 \\ \hline 0 & \cdots \quad 0 & -B_{j-1}^{(i)} & A_j^{(i)} \end{array}\right].$$

We shall need the following lemma.

LEMMA 3.3. *For each nonnegative integer n let A_n be a square matrix of order $n+1$ such that, if $n \geqq 1$,*

$$A_n = [\lambda_{ij}]_{0 \leqq i,j \leqq n} = \begin{bmatrix} A_{n-1} & \beta_n \\ \alpha_n & \lambda_{nn} \end{bmatrix} \tag{3.8}$$

where $\alpha_n = [\lambda_{n1}, \cdots, \lambda_{n,n-1}]$ and $\beta_n^t = [\lambda_{1n}, \cdots, \lambda_{n-1,n}]$. Suppose that $\det A_j > 0$, $j \geqq 0$, *and for each $n > 0$ we have*

$$\lambda_{nj} \leqq 0, \qquad 0 \leqq j \leqq n-1, \tag{3.9}$$

and

$$\lambda_{in} \leqq 0, \qquad 0 \leqq i \leqq n-1. \tag{3.10}$$

Then for each $n \geqq 0$, A_n^{-1} is nonnegative. If strict inequality holds for at least one j in (3.9) *and at least one i in* (3.10) *for each $n > 0$, then A_n^{-1} is positive.*

Proof. The lemma is obviously true for $n = 0$. Assume that A_{n-1}^{-1} is nonnegative (positive) and consider A_n^{-1}. Let

$$\sigma_n = \frac{\det A_{n-1}}{\det A_n} > 0.$$

Since A_n has the form (3.8), we may express A_n^{-1} as

$$A_n^{-1} = \begin{bmatrix} A_{n-1}^{-1} + \sigma_n (A_{n-1}^{-1}\beta_n)(\alpha_n A_{n-1}^{-1}) & -\sigma_n (A_{n-1}^{-1}\beta_n) \\ -\sigma_n(\alpha_n A_{n-1}^{-1}) & \sigma_n \end{bmatrix}.$$

From the induction assumption, we see that A_n^{-1} is nonnegative (positive) and the proof is complete.

Thus, Lemma 3.3 implies that the inverse of $\Gamma_j^{(i)}$ is nonnegative, $1 \leqq i \leqq k$, $j > i$. Hence, equation (3.7) implies that $\Phi_{nk}^{-1}(0)$ is nonnegative. Since the system of equations (3.3) is not diagonal if $c = 0$, we consider the following system:

$$\Phi_{nk}^{-1}(0) \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ \rho_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ \lambda_n \end{bmatrix} = \Phi_{nk}^{-1}(0)\Phi_{nk}(c) \begin{bmatrix} a_{kn}^{(k)} \\ a_{k+1,n}^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ a_{nn}^{(k)} \end{bmatrix}. \tag{3.11}$$

Since $\Phi_{nk}^{-1}(0)$ is lower triangular with positive diagonal entries, we see that $\lambda_n > 0$. The matrix $\Phi_{nk}^{-1}(0)\Phi_{nk}(c)$ has negative entries above the diagonal since this was true of $\Phi_{nk}(c)$. Also, the matrix $\Phi_{nk}^{-1}(0)\Phi_{nk}(c)$ has the recursive form (3.8) of Lemma 3.3. Moreover, since $\det \Phi_{nk}^{-1}(0)\Phi_{nk}(c) > 0$ for small positive c, a continuity argument shows it to be true for all positive c. We now prove that the elements below the diagonal of $\Phi_{nk}^{-1}(0)\Phi_{nk}(c)$ are negative.

LEMMA 3.4. *For $0 < c < 1$ the elements below the diagonal of the matrix $\Phi_{nk}^{-1}(0)\Phi_{nk}(c)$ are negative.*

Proof. Define

$$B_n(a) = \begin{bmatrix} \beta_k(a) \\ \cdot \\ \cdot \\ \cdot \\ \beta_n(a) \end{bmatrix} \tag{3.12}$$

where the $\{\beta_s(a)\}$ are as in Lemma 2.1, and again $a=\nu+c$. Recall from Lemma 2.1 that

$$h_{nk}(x;a)=x^{\nu}(x+\alpha)^c-\sum_{j=k}^{n}\beta_j(a)p_{jk}(x)$$

satisfies

$$\int_0^\infty h_{nk}(x;a)x^j\,du(x)=0, \qquad 0\leqq j\leqq n-k,$$

which may be written

$$\int_0^\infty h_{nk}(x;a)\phi_r(x)\,du(x)=0, \qquad 0\leqq r\leqq n-k, \tag{3.13}$$

$\phi_r(x)$ as in (3.1). Thus, (3.13) implies that

$$\sum_{j=k}^{n}\beta_j(a)\int_0^\infty p_{jk}(x)\phi_r(x)\,du(x)=\int_0^\infty x^{\nu}(x+\alpha)^c\phi_r(x)\,du(x), \qquad 0\leqq r\leqq n-k. \tag{3.14}$$

The system of equations (3.14) may be written

$$B_n(a)=\Phi_{nk}^{-1}(0)\begin{bmatrix}\langle\phi_0, x^{\nu}(x+\alpha)^c\rangle\\ \langle\phi_1, x^{\nu}(x+\alpha)^c\rangle\\ \cdot\\ \cdot\\ \cdot\\ \langle\Phi_{n-k}, x^{\nu}(x+\alpha)^c\rangle\end{bmatrix}. \tag{3.15}$$

From (3.4), the columns of $\Phi_{nk}^{-1}(0)\Phi_{nk}(c)$ may be written

$$\Phi_{nk}^{-1}(0)\begin{bmatrix}\int_0^\infty p_{sk}(x)\phi_0(x)(x+\alpha)^c\,du(x)\\ \int_0^\infty p_{sk}(x)\phi_1(x)(x+\alpha)^c\,du(x)\\ \cdot\\ \cdot\\ \cdot\\ \int_0^\infty p_{sk}(x)\phi_{n-k}(x)(x+\alpha)^c\,du(x)\end{bmatrix}\equiv A_s, \qquad k\leqq s\leqq n.$$

That is,

$$\Phi_{nk}^{-1}(0)\Phi_{nk}(c)=[A_k, A_{k+1},\cdots,A_n].$$

But (1.18) and (3.15) imply that

$$A_k=B_n(c), \tag{3.16a}$$

$$A_s=B_n(c)-\sum_{j=k+1}^{s}b_{js}^{(k)}B_n(j+c), \qquad k+1\leqq s\leqq n. \tag{3.16b}$$

Let

$$A_s=\begin{bmatrix}\varepsilon_{ks}\\ \varepsilon_{k+1,s}\\ \cdot\\ \cdot\\ \cdot\\ \varepsilon_{ns}\end{bmatrix};$$

in particular

$$A_k = B_n(c) = \begin{bmatrix} \varepsilon_{kk} \\ \varepsilon_{k+1,k} \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_{nk} \end{bmatrix}.$$

Since $p_{r,0}(x) = \gamma_r \phi_r(x)$ $(\gamma_r > 0)$, Theorem 2.2 implies that

$$\int_0^\infty \phi_r(x)(x+\alpha)^c\, du(x) < 0, \qquad r > 0.$$

Thus, recalling the characteristics of $\Phi_{nk}^{-1}(0)$ we see that $\varepsilon_{jk} < 0$, $k+1 \leqq j \leqq n$. From Lemma 2.1 and the inequality (1.19) $b_{js}^{(k)}\beta_{j+r}(j+c) > 0$, $1 \leqq r \leqq n-j$, for each j, $k+1 \leqq j \leqq s \leqq n-1$. Recalling (3.16) we see that $\varepsilon_{rs} < 0$, $k \leqq s < r \leqq n$. This completes the proof of the lemma.

It now follows from Lemma 3.3 that the matrix $\Phi_{nk}^{-1}(0)\Phi_{nk}(c)$ has a positive inverse. Recalling the system of equations (3.11) we see that, for $0 < c < 1$ the inequalities (1.15) are obtained. By iteration of this result Theorem 1.1 is proved.

4. An application of Theorem 1.1. At this point, recall (1.11). We see that by setting

$$du(x) = \begin{cases} F(x)\, dx, & 0 \leqq x \leqq 1, \\ 0, & x > 1, \end{cases}$$

and $dv(x) = x^{c_1}[1+\lambda x]^{c_2}\, du(x)$, we have the following result:

THEOREM 4.1. *Suppose $MV(n, k; \Phi)$ is stable for all $n \geqq k+1 \geqq 1$. Let*

$$\Gamma(\theta) = \left[\sin^2\left(\frac{\theta}{2}\right)\right]^{c_1}\left[1+\lambda \sin^2\left(\frac{\theta}{2}\right)\right]^{c_2}\Phi(\theta)$$

where c_1, c_2, λ are nonnegative numbers. Then $MV(n, k; \Gamma)$ is stable for all $n \geqq k+1 \geqq 1$.

Note that if we set $Q(x) = x^c \prod_{i=1}^r [1+\lambda_i x]^{c_i}$, $c_i \geqq 0$, $\lambda_i \geqq 0$, $1 \leqq i \leqq r$, and $\Delta(\theta) = Q(\sin^2(\theta/2))\Phi(\theta)$, then iteration of Theorem 4.1 yields stability of $MV(n, k; \Delta)$ for all $n \geqq k+1 \geqq 1$.

Acknowledgment. This work is an extension of a portion of the author's doctoral dissertation at Drexel University, Philadelphia, Pennsylvania, under the supervision of Professor William F. Trench. The author wishes to express his gratitude and appreciation to Professor Trench for his encouragement and helpful suggestions.

REFERENCES

[1] R. ASKEY, *Orthogonal expansions with positive coefficients*, Proc. Amer. Math. Soc., 26 (1965), pp. 1191–1194.

[2] T. N. E. GREVILLE, *On a problem of E. L. DeForest in iterated smoothing*, this Journal, 5 (1974), pp. 376–397.

[3] ———, *On the stability of linear smoothing formulas*, SIAM J. Numer. Anal. 3 (1966), pp. 157–170.

[4] M. RIESZ, *Sur le probleme des moments, Troisième Note*, Arkiv för Matematik, Astronomi och Fysik, 17 (16) (1923), 52 pp.

[5] I. J. SCHOENBERG, *Some analytical aspects of the problem of smoothing*, Studies and Essays presented to R. Courant on His 60th Birthday, Interscience, New York, 1948, pp. 351–370.

[6] G. SZEGÖ, *Orthogonal Polynomials*, rev. ed., AMS Colloquium Publications, vol. 23, American Mathematical Society, Providence, RI, 1959.

[7] W. F. TRENCH, *Nonnegative and alternating expansions of one set of orthogonal polynomials in terms of another*, this Journal, 4 (1973), pp. 111–115.

[8] ———, *Discrete minimum variance smoothing of a polynomial plus random noise*, J. Math. Anal. Appl., 35 (1971), pp. 630–645.

[9] ———, *Orthogonal polynomial expansions with nonnegative coefficients*, this Journal, 7 (1976), pp. 824–833.

[10] ———, *Proof of a conjecture of Askey on orthogonal expansions with positive coefficients*, Bull. Amer. Math. Soc., 81 (1975), pp. 954–956.

[11] ———, *Stability of a class of discrete minimum variance smoothing formulas*, SIAM J. Numer. Anal., 9 (1972), pp. 307–315.

[12] ———, *Strong discrete minimum variance estimation of a polynomial plus random noise*, J. Math. Anal. Appl., 41, (1973), pp. 20–33.

SIAM J. MATH. ANAL.
Vol. 9, No. 3, June 1978

Copyright © 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0903-0007$01.00/0

AN ITERATIVE TECHNIQUE FOR OBTAINING SOLUTIONS OF A THOMAS-FERMI EQUATION*

C. D. LUNING†

Abstract. A convergent iteration scheme is given for the boundary value problem $y''=x^{-1/2}y^{3/2}$, $y(0)=1$, $-y(b)+by'(b)=0$. An iteration scheme based on eigenpairs of Hilbert–Schmidt operators obtained from a Green's function representation for solutions of a linearization of the eigenvalue problem $u''=\lambda x^{-1/2}u^{3/2}$, $-\alpha u(0)+u'(0)=0$, $-u(1)+u'(1)=0$ is shown to converge to a solution. The solution of the Thomas–Fermi equation for $b=\lambda^{2/3}$ is then given by $y(x)=u(\lambda^{-2/3}x)$. Since the linear operators involved are Hilbert–Schmidt, the iteration lends itself to implementation through Galerkin methods.

1. Introduction. In 1927, in their work on potentials and charge densities in atoms, L. H. Thomas [22] and E. Fermi [7] independently introduced the nonlinear second order differential equation

$$y''=x^{-1/2}y^{3/2} \tag{1.1}$$

known as the Thomas–Fermi equation. Three sets of boundary values, corresponding to three different physical situations, are commonly associated with (1.1):

a) the neutral atom with Bohr radius b

$$y(0)=1, \qquad -y(b)+by'(b)=0, \tag{1.2}$$

b) the isolated neutral atom

$$y(0)=1, \qquad \lim_{x\to\infty} y(x)=0, \tag{1.3}$$

c) the ionized atom

$$y(0)=1, \qquad y(a)=0. \tag{1.4}$$

We note that the Thomas–Fermi equations are still used for atomic calculations (see for example [4], [5], [21]), and the related Thomas–Fermi theory is still a subject of physical research (see [13], [14]). In this paper we restrict our considerations to equation (1.1) subject to the boundary values (1.2). We are not concerned here with the physical ramifications of the Thomas–Fermi theory, but only with some of the mathematical aspects of showing that a certain sequence of functions does converge to a solution. In fact it is shown that for certain values of b, a sequence of functions obtained from the solutions of a related linear eigenvalue problem converges uniformly on $[0, b]$ to the solution of the Thomas–Fermi equation with boundary values (1.2).

Before describing the approach used in this paper, a short synopsis of other approaches which have been used is given. The first boundary values to be considered were those of the isolated neutral atom (1.3). For this case the existence of a unique solution follows from a general theorem of A. Mambriani [15]. For computational purposes, Thomas utilized Adam's method of numerical integration of the differential equation. Sommerfeld [20] developed the approximate solution

$$y(x)=144x^{-3}[1+(144x^{-3})^{m_1/3}]^{m_2/3} \tag{1.5}$$

* Received by the editors December 11, 1975, and in revised form October 7, 1976.
† Department of Mathematics, Texas A & M University, College Station, Texas 77843.

where $m_2 < 0 < m_1$ are the roots of $m^2 + 7m - 6 = 0$. Sommerfeld's approximation is quite accurate for large x but underestimates the solution near the origin (see [6]). Analogue computers have also been used to find approximate solutions (see [1]), and more recently Ramnath [18] has used a technique known as multiple scales.

Passing now to the other boundary conditions, solution by infinite series has been used in conjunction with all three sets of boundary values (1.2), (1.3) and (1.4). Equation (1.1) with the initial condition $y(0) = 1$ leads to the formal series solution

$$y(x) = 1 + b_2 x + b_3 x^{3/2} + \cdots + b_k x^{k/2} + \cdots \tag{1.6}$$

where the coefficients b_k, $k \geqq 3$, can be expressed as polynomials in the coefficient b_2. Of course b_2 is the slope of y at the origin. There is a critical value ω, ω approximately -1.588, such that if $b_2 > \omega$, then the series converges to a solution satisfying the boundary values (1.2); if $b_2 = \omega$, then the series converges to the solution satisfying the boundary values (1.3); and finally if $b_2 < \omega$, then the series converges to a solution satisfying the boundary values (1.4). E. Hille [9], [10] has a discussion of the convergence of the series (1.6). These papers also contain references to the numerous numerical results which have been obtained from the series solution.

We mention that there has been an abundance of literature on the Thomas–Fermi equation and related theories. Extensive reviews of this literature can be found in [8], [16], [23].

The approach taken in this paper is decidedly different from those previously described. As indicated before we restrict our attention to

$$\begin{aligned} &y'' = x^{-1/2} y^{3/2}, \\ &y(0) = 1, \qquad -y(b) + by'(b) = 0. \end{aligned} \tag{1.7}$$

An iterative procedure is used to generate a sequence of functions $\{y_k\}$ which is shown to converge to a solution of (1.7). The iteration is based upon the eigenvalues and eigenfunctions of a related linear Sturm–Liouville problem. The proof is valid however only for a limited range of values for the end point b.

We first consider the nonlinear eigenvalue problem

$$\begin{aligned} &u'' = \lambda x^{-1/2} u^{3/2}, \\ &u(0) = 1, \quad u'(0) = \alpha, \quad -u(1) + u'(1) = 0. \end{aligned} \tag{1.8}$$

If u is a positive solution of (1.8) with corresponding positive eigenvalue λ, then $y(x) = u(\lambda^{-2/3} x)$ is a solution of (1.7) for $b = \lambda^{2/3}$. The main result can be stated as

THEOREM 1. *For $\alpha \geqq -1$, let $u_0(x) = 1 + \alpha x$, and $f_0 = u_0^{1/2}$. For each $k = 1, 2, \cdots$, let (u_k, λ_k) denote the positive solution of*

$$\begin{aligned} &u'' = \lambda x^{-1/2} f_{k-1}(x) u, \qquad 0 < x < 1, \\ &-\alpha u(0) + u'(0) = 1, \qquad -u(1) + u'(1) = 0 \end{aligned} \tag{1.9}$$

where u_k is normalized by $u_k(0) = 1$ and where $f_k = u_k^{1/2}$. Then $u_0(x) < u_k(x)$, $0 < x < 1$, $\lambda_{2k} < \lambda_{2k+2} < \lambda_{2k+1} < \lambda_{2k-1}$, and $u_{2k}(x) < u_{2k+2}(x) < u_{2k+1}(x) < u_{2k-1}(x)$, $0 < x < 1$. Moreover, there is a positive solution (u, λ) of (1.8) such that $\lambda_k \to \lambda$ and $u_k \to u$ uniformly on $[0, 1]$.

As for the restriction on α, in § 2, equation (1.9) is transformed into an equivalent integral equation which has a pointwise positive function as its kernel when $\alpha \geqq -1$. This then insures the existence of the required positive eigenvalues and eigenfunctions.

We note that Moore and Nehari [17] have used a similar iteration in their consideration of the nonlinear boundary value problem

$$(1.10)\qquad \begin{aligned} &y''+p(x)y^{2n+1}=0, \qquad a<x<b,\\ &y(a)=0, \qquad y'(b)=0 \end{aligned}$$

where p is positive and continuous on $(0, \infty)$ and n is a positive integer. Their approach is also based upon a sequence of eigenvalues and eigenfunctions of a related linear Sturm–Liouville problem. Briefly they proceed as follows: Let $\{u_\nu\}_{\nu=1}^{\infty}$ be a minimizing sequence of admissible functions for the generalized Rayleigh quotient

$$(1.11)\qquad J(y)=\frac{(\int_a^b (y'(x))^2\,dx)^{n+1}}{\int_a^b p(x)y^{2n+2}(x)\,dx}.$$

For each ν, iteratively define the sequence $\{v_{\nu,\mu}\}_{\mu=1}^{\infty}$: $v_{\nu,1}=u_\nu$. For $\mu \geqq 2$, $\alpha_{\nu,\mu}$ is the least eigenvalue of the system

$$(1.12)\qquad \begin{aligned} &v''+\lambda p(x)v_{\nu,\mu-1}^{2n}v=0, \qquad a<x<b,\\ &v(a)=0, \qquad v'(b)=0 \end{aligned}$$

and $v_{\nu,\mu}$ is the corresponding positive eigenfunction normalized by $\int_a^b (v'_{\nu,\mu}(x))^2\,dx=1$. It is then shown that a subsequence of the diagonal sequence $\{(v_{\nu,\nu}, \alpha_{\nu,\nu})\}$ converges uniformly to a positive solution (v, λ) of

$$(1.13)\qquad \begin{aligned} &v''+\lambda p(x)v^{2n+1}=0,\\ &v(a)=0, \qquad v'(b)=0. \end{aligned}$$

A solution of (1.10) is then obtained from the solution of (1.13) by $y=\lambda^{1/2n}v$.

It appears possible to modify the Moore–Nehari approach to accommodate the Thomas–Fermi equation. However, the Moore–Nehari convergence is of a subsequential nature whereas with our approach it is shown that the original sequence converges.

As for the remainder of this paper, in §2 some preliminary structure concerning the related linear Sturm–Liouville problem is developed. In § 3 the main result is proved. In § 4 some uniqueness results are given. Also the dependence of the end point b on the initial slope α is discussed. Finally mention is made of how the iteration might be implemented through Galerkin methods.

2. Preliminaries. The form of the linear Sturm–Liouville problem used in the iteration is

$$(2.1)\qquad \begin{aligned} &u''-\lambda x^{-1/2}f(x)u=0,\\ &-\alpha u(0)+u'(0)=0, \qquad -u(1)+u'(1)=0 \end{aligned}$$

where f is continuous on $[0, 1]$, $f(0)=1$, $f(x)\geqq 1-x$, $0\leqq x\leqq 1$. By using a Green's function, (2.1) can be written as the equivalent integral equation:

$$(2.2)\qquad \begin{aligned} &u(x)=\int_0^1 K(x, t; \alpha)t^{-1/2}f(t)u(t)\,dt,\\ &K(x, t; \alpha)=\begin{cases} x(1+\alpha t), & 0\leqq t\leqq x,\\ (1+\alpha x)t, & x\leqq t\leqq 1. \end{cases} \end{aligned}$$

Let H_f denote the Hilbert space:

$$H_f = \left\{ u : \int_0^1 x^{-1/2} f(x) |u(x)|^2 \, dx < \infty \right\}, \tag{2.3}$$
$$\|u\|_f = \left(\int_0^1 x^{-1/2} f(x) |u(x)|^2 \, dx \right)^{1/2}.$$

On H_f define the linear integral operator T_f by

$$(T_f u)(x) = \int_0^1 K(x, t; \alpha) t^{-1/2} f(t) u(t) \, dt \tag{2.4}$$

and on $L^2[0, 1]$ define the linear integral operator M_f by

$$(M_f \phi)(x) = \int_0^1 x^{-1/4} f^{1/2}(x) K(x, t; \alpha) t^{-1/4} f^{1/2}(t) \phi(t) \, dt. \tag{2.5}$$

Since M_f has a square integrable symmetric kernel, M_f is a self-adjoint Hilbert–Schmidt operator on $L^2[0, 1]$ (see [19]). Since for $\alpha \geqq -1$, the kernel of M_f is a positive function, it follows that M_f has at least one positive eigenvalue and that there is a positive eigenfunction associated with the largest positive eigenvalue (see [11]). For the remainder of this paper it will be assumed that $\alpha \geqq -1$.

The spectral properties of T_f necessary for our computations are readily obtainable from those of M_f by noting that if $\phi \in L^2[0, 1]$ is an eigenfunction of M_f with corresponding eigenvalue ν, then for the function u in H_f defined by

$$u = x^{1/4} f^{-1/2}(x) \phi \tag{2.6}$$

it follows from (2.4) and (2.5) that

$$x^{-1/4} f^{1/2}(x) T_f(u) = M_f(\phi) = \nu \phi = \nu x^{-1/4} f^{1/2}(x) u. \tag{2.7}$$

Thus T_f and M_f have the same spectrum and their eigenfunctions are related by (2.6). By combining these spectral properties with the definition (2.3) of the inner product in H_f, we conclude that T_f is a selfadjoint Hilbert–Schmidt operator in H_f with at least one positive eigenvalue. Associated with the largest positive eigenvalue is a positive eigenfunction.

Since T_f is a selfadjoint Hilbert–Schmidt operator, it has a complete orthogonal system of eigenfunctions in H_f. The smoothness of the kernel of T_f allows us to assume that the eigenfunctions are $C^1[0, 1]$. If u is a $C^1[0, 1]$ eigenfunction of T_f, that is if $T_f u = \nu u$ then

$$u'' - \nu^{-1} x^{-1/2} f(x) u = 0, \tag{2.8}$$
$$-\alpha u(0) + u'(0) = 0, \qquad -u(1) + u'(1) = 0.$$

Thus $u(0) \neq 0$ for otherwise $u'(0) = 0$ also which would imply $u(x) \equiv 0$, $0 \leqq x \leqq 1$ (see [2]). We normalize the eigenfunctions of T_f by $u(0) = 1$. Of course by the orthogonality of the eigenfunctions, there can be at most one positive $C^1[0, 1]$ eigenfunction of T_f.

We now prove a technical lemma to be used in the convergence proof. Let h_1 and h_2 be continuous positive functions on $[0, 1]$ satisfying

i) there is a $\delta>0$ such that $h_1(x)<h_2(x)$, $0\leqq x<\delta$,

ii) if there is an $x_0\in(0,1)$ such that $h_1(x_0)>h_2(x_0)$, then $h_1(x)>h_2(x)$, $x\in[x_0,1]$.

LEMMA 1. *If h_1 and h_2 are as above and if for $i=1,2$, v_i is a positive continuous solution of*

$$y''=x^{-1/2}h_i(x)y, \qquad 0<x<1, \tag{2.9}$$
$$y(0)=1, \quad y'(0)=\alpha, \quad -y(1)+y'(1)=0$$

then $v_1(x)<v_2(x)$, $0<x<1$.

Proof. By the assumptions on h_1 and h_2 we have $h_1(0)v_1(0)<h_2(0)v_2(0)$. From the continuity it follows that there is an $\varepsilon>0$ such that for $0<x<\varepsilon$, $v_1''(x)<v_2''(x)$. Utilizing the initial values, we have by integration that $v_1'(x)<v_2'(x)$, $v_1(x)<v_2(x)$, $(0<x<\varepsilon$. Assume $x_1\in(0,1)$ exists such that $v_1(x_1)=v_2(x_1)$ and $v_1(x)<v_2(x)$, $0<x<x_1$. Then $v_1'(x_1)\geqq v_2'(x_1)$ and since $v_1'(x)<v_2'(x)$, $0<x<\varepsilon$, there must exist $x_2\in(0,x_1)$ such that $v_1''(x_2)>v_2''(x_2)$. That is, $x_2^{-1/2}h_1(x_2)v_1(x_2)>x_2^{-1/2}h_2(x_2)v_2(x_2)$. Since $v_1(x_2)<v_2(x_2)$ it follows that $h_1(x_2)>h_2(x_2)$. Thus by our assumptions, $h_1(x)>h_2(x)$, $x_2\leqq x\leqq 1$. Using (2.9), we immediately obtain $v_1''(x)>v_2''(x)$, $v_1'(x)>v_2'(x)$ and $v_1(x)>v_2(x)$, $x_1<x\leqq 1$. Let $g_1(x)=-v_1(x)+xv_1'(x)$ and $g_2(x)=-v_2(x)+xv_2'(x)$. Then $g_1(x_1)\geqq g_2(x_1)$ and $g_1'(x)>g_2'(x)$, $x_1\leqq x\leqq 1$. Thus $g_1(1)>g_2(1)$. However from the boundary values, $g_1(1)=g_2(1)=0$. Thus x_1 does not exist and $v_1(x)<v_2(x)$, $0<x<1$.

We conclude this section with an observation on the behavior of the eigenvalues of equation (2.1).

LEMMA 2. *If f_1 and f_2 are two distinct functions satisfying the conditions of f in* (2.1) *such that $f_1(x)\leqq f_2(x)$ and if (w_1,μ_1) and (w_2,μ_2) are the positive normalized solutions corresponding respectively to f_1 and f_2, then $\mu_2<\mu_1$.*

The proof of Lemma 2 is omitted since Lemma 2 is essentially Sturm's second comparison theorem, the proof of which can be found in many standard books (see for example [3]).

3. The convergence of the iteration. We now prove Theorem 1. Let $u_0, f_0, \{u_k\}_{k=1}^{\infty}$, $\{\lambda_k\}_{k=1}^{\infty}$ and $\{f_k\}_{k=1}^{\infty}$ be as in Theorem 1. Since $u_k''(x)>0$ and $u_k'(0)=\alpha$, $u_k(0)=1$, it follows that for each $k\geqq 1$, $u_k(x)>1+\alpha x=u_0(x)$, $0<x\leqq 1$.

LEMMA 3. *$\lambda_2<\lambda_1$, $\lambda_2 f_1$ can intersect $\lambda_1 f_0$ at most once in* $(0,1)$, *and $u_2(x)<u_1(x)$,* $0<x<1$.

Proof. $\lambda_2<\lambda_1$ follows from Lemma 2 and the fact that $f_0(x)<f_1(x)$, $0<x\leqq 1$. Since $\lambda_2^2u_1(0)<\lambda_1^2u_0(0)$ and since $\lambda_2^2u_1'(x)$ is increasing and $\lambda_1^2u_0'(x)$ is constant, it follows that $\lambda_2^2u_1(x)$ can intersect $\lambda_1^2u_0(x)$ at most once in $(0,1)$. Thus $\lambda_2 f_1$ can intersect $\lambda_1 f_0$ at most once in $(0,1)$. We can now apply Lemma 1 with $h_1=\lambda_2 f_1$ and $h_2=\lambda_1 f_0$ to conclude $u_2(x)<u_1(x)$, $0<x<1$.

LEMMA 4. *$\lambda_2<\lambda_3<\lambda_1$, $\lambda_2 f_1$ can intersect $\lambda_3 f_2$ at most once in* $(0,1)$, *$\lambda_3 f_2$ can intersect $\lambda_1 f_0$ at most once in* $(0,1)$, *and $u_2(x)<u_3(x)<u_1(x)$, $0<x<1$.*

Proof. $\lambda_2<\lambda_3<\lambda_1$ follows from Lemma 2 and the fact that $f_0(x)<f_2(x)<f_1(x)$, $0<x<1$. Since $f_1(0)=f_2(0)=1$, we have $\lambda_2 f_1(0)<\lambda_3 f_2(0)$. As soon as it is shown that $\lambda_2 f_1(x)$ can cross $\lambda_3 f_2(x)$ at most once in $(0,1)$, it follows from Lemma 1 that $u_2(x)<u_3(x)$, $0<x<1$. Suppose there exist a_0 and a_1, $0<a_0<a_1<1$ such that $\lambda_2 f_1(a_0)=\lambda_3 f_2(a_0)$, $\lambda_2 f_1(a_1)=\lambda_3 f_2(a_1)$, $\lambda_2 f_1(x)\leqq\lambda_3 f_2(x)$, $0\leqq x\leqq a_0$, $\lambda_2 f_1(x)>$ $\lambda_2 f_1(x)>\lambda_3 f_2(x)$, $a_0<x<a_1$. Let

$$z_1=\lambda_2^2 f_1^2=\lambda_2^2 u_1 \quad\text{and}\quad z_2=\lambda_3^2 f_2^2=\lambda_3^2 u_2. \tag{3.1}$$

Then

$$(3.2)\quad \begin{aligned} &z_1''(x)=\lambda_1 x^{-1/2}f_0(x)z_1(x), \qquad 0<x<1,\\ &z_2''(x)=\lambda_2 x^{-1/2}f_1(x)z_2(x), \qquad 0<x<1,\\ &z_1(a_0)=z_2(a_0), \qquad z_1(a_1)=z_2(a_1),\\ &z_1'(a_0)\geqq z_2'(a_0), \qquad z_1'(a_1)\leqq z_2'(a_1),\\ &z_1(x)>z_2(x), \qquad a_0<x<a_1,\\ &-z_1(1)+z_1'(1)=-z_2(1)+z_2'(1)=0. \end{aligned}$$

Thus there is an $x_3\in(a_0, a_1)$ such that $z_2''(x_3)>z_1''(x_3)$. By using the differential equations of (3.2) which z_1 and z_2 satisfy along with the fact that $z_1(x_3)>z_2(x_3)$, we conclude that $\lambda_2 f_1(x_3)>\lambda_1 f_0(x_3)$. From Lemma 3, $\lambda_2 f_1$ and $\lambda_1 f_0$ intersect at most once in (0, 1), thus $\lambda_2 f_1(x)>\lambda_1 f_0(x)$, $x_3<x\leqq 1$. Putting this in the differential equation satisfied by z_1 and z_2 we conclude $z_2(x)>z_1(x)$, $a_1<x\leqq 1$ and thus $z_2''(x)>z_1''(x)$, $a_1<x\leqq 1$. Let $w_1(x)=-z_1(x)+xz_1'(x)$, $w_2(x)=-z_2(x)+xz_2'(x)$. Then $w_2(a_1)\geqq w_1(a_1)$ and $w_2'(x)>w_1'(x)$, $a_1<x\leqq 1$. Thus $w_2(1)>w_1(1)$, which contradicts $w_2(1)=w_1(1)=0$. We thus have the desired result that $\lambda_2 f_1$ can cross $\lambda_3 f_2$ at most once in (0, 1).

The proof that $\lambda_3 f_2$ can intersect $\lambda_1 f_0$ at most once in (0, 1) and that $u_3(x)<u_1(x)$, $0<x<1$ is very similar to the proof of Lemma 3 and is omitted.

LEMMA 5. *For each* $k\geqq 1$, $\lambda_{2k}<\lambda_{2k+2}<\lambda_{2k+1}<\lambda_{2k-1}$, $u_{2k}(x)<u_{2k+2}(x)<u_{2k+1}(x)<u_{2k-1}(x)$, $0<x<1$.

The proof of Lemma 5 is omitted since it is accomplished inductively using essentially the same arguments as in the proof of Lemma 4.

The sequences $\{\lambda_{2k}\}_{k=1}^{\infty}$ and $\{u_{2k}\}_{k=1}^{\infty}$ are increasing bounded sequences whereas the sequences $\{\lambda_{2k+1}\}_{k=0}^{\infty}$ and $\{u_{2k+1}\}_{k=0}^{\infty}$ are decreasing bounded sequences all of which thus have limits. As $k\to\infty$, let $\lambda_*=\lim\lambda_{2k}$, $u_*=\lim u_{2k}$, $\lambda^*=\lim\lambda_{2k+1}$ and $u^*=\lim u_{2k+1}$. From (2.2) we have

$$(3.3)\quad u_n=\lambda_n\int_0^1 K(x,t;\alpha)t^{-1/2}u_{n-1}^{1/2}(t)u_n(t)\,dt.$$

Utilizing the dominated convergence theorem, it follows that

$$(3.4)\quad \begin{aligned} u^*&=\lambda^*\int_0^1 K(x,t;\alpha)t^{-1/2}u_*^{1/2}(t)u^*(t)\,dt,\\ u_*&=\lambda_*\int_0^1 K(x,t;\alpha)t^{-1/2}u^{*1/2}(t)u_*(t)\,dt. \end{aligned}$$

Of course $u_*(0)=u^*(0)=1$, $u_*(x)\leqq u^*(x)$, $0\leqq x\leqq 1$, and $\lambda_*\leqq\lambda^*$. From (3.4) it follows that

$$(3.5)\quad \begin{aligned} u^{*\prime\prime}&=\lambda^*x^{-1/2}u_*^{1/2}u^*,\\ u_*''&=\lambda_*x^{-1/2}u^{*1/2}u_*, \end{aligned}\qquad 0<x<1,$$

and that both u_* and u^* satisfy the boundary conditions

$$(3.6)\quad -\alpha u(0)+u'(0)=0, \qquad -u(1)+u'(1)=0.$$

LEMMA 6. $\lambda_*=\lambda^*$, $u_*(x)=u^*(x)$, $0\leqq x\leqq 1$.

Proof. If $u_*(x)\neq u^*(x)$ then by Lemma 2, $\lambda_*<\lambda^*$. From the equations (3.5) and the initial values it follows that there is a $\delta>0$ such that $u_*''(x)<u^{*\prime\prime}(x)$, $u_*'(x)<u^{*\prime}(x)$ and

$u_*(x) < u^*(x)$, $0 < x < \delta$. If there is an $x_0 \in (0, 1)$ such that $u_*(x_0) = u^*(x_0)$, $u_*(x) < u^*(x)$, $0 < x < x_0$, then there is an $x_1 \in (0, x_0)$ such that $u_*''(x_1) > u^{*\prime\prime}(x_1)$, that is

$$\lambda_* x_1^{-1/2} u^{*1/2}(x_1) u_*(x_1) > \lambda^* x_1^{-1/2} u_*^{1/2}(x_1) u^*(x_1)$$

from which it follows that $\lambda_* u_*^{1/2}(x_1) > \lambda^* u^{*1/2}(x_1)$, a contradiction. Thus if $u_*(x) \neq u^*(x)$ then $u_*(x) \lneq u^*(x)$, $0 < x < 1$. By utilizing this along with $\lambda_* < \lambda^*$ in (3.5), we conclude $u_*''(x) < u^{*\prime\prime}(x)$, $0 < x < 1$. Let $g_*(x) = -u_*(x) + x u_*'(x)$, $g^*(x) = -u^*(x) + x u^{*\prime}(x)$. Then $g_*(0) = g^*(0) = -1$ and $g_*'(x) < g^{*\prime}(x)$, $0 < x < 1$. Thus $g_*(1) < g^*(1)$. However from the boundary conditions it follows that $g_*(1) = g^*(1) = 0$. Thus $u_*(x) = u^*(x)$, $0 \leqq x \leqq 1$ and then of course $\lambda_* = \lambda^*$.

We have shown that the sequence $\{(u_k, \lambda_k)\}_{k=1}^{\infty}$ converges to a positive solution of (1.8). To complete the proof of Theorem 1, we need only show that the convergence of $\{u_k\}$ is uniform. The sequence of functions $\{u_k\}$ is uniformly bounded on $[0, 1]$ by u_0 and u_1. From (3.3), it can be shown that the sequence $\{u_k\}$ is equicontinuous on $[0, 1]$ and thus the convergence of u_k to u is uniform on $[0, 1]$.

4. Uniqueness results and eigenvalue bounds. In this section we show that the positive $C^1[0, 1]$ solution of (1.8) is unique. We first consider the initial value problem

$$\text{(4.1)} \qquad \begin{aligned} &y'' = \beta x^{-1/2} y^{3/2}, \\ &y(0) = 1, \qquad y'(0) = \alpha, \end{aligned}$$

for fixed $\beta > 0$. If y_1 and y_2 are two positive continuous solutions of (4.1) and if there is an $x_0 > 0$ such that $y_1(x_0) > y_2(x_0)$, then for $x_1 = \sup\{x < x_0; y_1(x) = y_2(x)\}$, it is clear from (4.1) that $y_1(x) = y_2(x)$, $0 \leqq x \leqq x_1$ and $y_1''(x) > y_2''(x)$, $y_1'(x) > y_2'(x)$, $y_1(x) > y_2(x)$, $x > x_1$.

LEMMA 7. *The* $C^1[0, 1]$ *solution of* (4.1) *is unique.*

Proof. If y_1 and y_2 are two distinct $C^1[0, 1]$ solutions of (4.1), let x_1 be as above. We show x_1 does not exist. Choose $x_2 > x_1$ so that $y_1(x_2)$, $y_1'(x_2)$, $y_2(x_2)$, $y_2'(x_2)$ all exist. Then for $0 < x < x_2$

$$\begin{aligned} 0 &\leqq y_1'(x) - y_2'(x) + \beta x^{-1/2}(y_1^{3/2}(x) - y_2^{3/2}(x)) \\ &\leqq (x_2^{1/2} + \beta) x^{-1/2}(y_1'(x) - y_2'(x) + 2 y_1^{1/2}(x)(y_1(x) - y_2(x))). \end{aligned}$$

Since y_1 is continuous on $[0, x_2]$, there is $k > 0$ such that

$$\begin{aligned} 0 &\leqq y_1'(x) - y_2'(x) + \beta x^{-1/2}(y_1^{3/2}(x)) \\ &\leqq k x^{-1/2}(y_1(x) - y_2(x) + y_1'(x) - y_2'(x)), \end{aligned}$$

$0 < x < x_2$. Let $P(x) = y_1(x) - y_2(x) + y_1'(x) - y_2'(x)$, $\rho(x) = P(x_2) \exp(2k(x^{1/2} - x_2^{1/2}))$. Then $\rho(x)$ satisfies $\rho' = k x^{-1/2} \rho$, $0 < x < x_2$, $\rho(x_2) = P(x_2) > 0$. Moreover $\rho(x) \leqq P(x)$, $0 < x \leqq x_2$ for otherwise there are t and h such that $0 < t - h < t \leqq x_2$, $\rho(t) = P(t)$ and $\rho(x) > P(x)$, $t - h \leqq x < t$. However

$$\begin{aligned} P(t) - P(t-h) &= \int_{t-h}^{t} P'(x)\,dx = \int_{t-h}^{t} y_1'(x) - y_2'(x) + \beta x^{-1/2}(y_1^{3/2}(x) - y_2^{3/2}(x))\,dx \\ &\leqq k \int_{t-h}^{t} x^{-1/2} P(x)\,dx \leqq k \int_{t-h}^{t} x^{-1/2} \rho(x)\,dx = \int_{t-h}^{t} \rho'(x)\,dx \\ &= \rho(t) - \rho(t-h) = P(t) - \rho(t-h). \end{aligned}$$

Thus $\rho(t-h) \leqq P(t-h)$, a contradiction. However $\rho(x) \leqq P(x)$, $0 < x \leqq x_2$ is obviously a contradiction since

$$\lim_{x\to 0^+} \rho(x) = P(x_2) \exp(-2kx_2^{1/2}) > 0$$

and $\lim_{x\to 0^+} P(x) = 0$. Thus $y_1 = y_2$.

THEOREM 2. *The positive* $C^1[0, 1]$ *solution* (u, λ) *of* (1.8) *is unique.*

Proof. Suppose (u, λ) and $(\hat{u}, \hat{\lambda})$ are two positive $C^1[0, 1]$ solutions of (1.8). If $\lambda \neq \hat{\lambda}$ then we can proceed as in the proof of Lemma 6 to arrive at a contradiction. Thus $\lambda = \hat{\lambda}$ and by Lemma 7, $u = \hat{u}$.

We now give some bounds for the eigenvalues λ of (1.8), and describe the behavior of the end point b as α goes from -1 to $+\infty$. First we show that increasing α decreases λ and of course decreases b.

THEOREM 3. *If* $\alpha_1 > \alpha_2 \geqq -1$ *and if* (u_1, λ_1) *and* (u_2, λ_2) *are respectively the positive solutions of* (1.8) *with initial slopes* $u_1'(0) = \alpha_1$ *and* $u_2'(0) = \alpha_2$ *then* $\lambda_1 < \lambda_2$.

Proof. If $\lambda_1 \geqq \lambda_2$ then, since $u_1(0) = u_2(0)$, $u_1'(0) > u_2'(0)$ and $u_1''(x) = \lambda_1 x^{-1/2} u_1^{3/2}(x)$, $u_2''(x) = \lambda_2 x^{-1/2} u_2^{3/2}(x)$, it follows that $u_1(x) > u_2(x)$ and $u_1''(x) > u_2''(x)$, $0 < x \leqq 1$. Thus $-u_1(x) + xu_1'(x)$ and $-u_2(x) + xu_2'(x)$ can not both be zero at $x = 1$, contradicting the boundary condition.

By combining Theorem 3 with the result of Theorem 1 that $\lambda(\alpha) < \lambda_1(\alpha)$ we obtain the estimate that for $\alpha \geqq -1$

$$\lambda^{-1}(\alpha) \geqq \lambda^{-1}(\alpha = -1) > \lambda_1^{-1}(\alpha = -1) = \sup_{\phi \in H_{f_0}} \frac{(T_{f_0}\phi, \phi)_{f_0}}{(\phi, \phi)_{f_0}} \tag{4.2}$$

where $f_0 = (1-x)^{1/2}$. By letting $\phi(x) = 1$ and by performing the integrations in (4.2) we obtain the bound $\lambda < 6\sqrt{6}/5$. For the case $\alpha \geqq 0$, we have $f_0(x) = (1 + \alpha x)^{1/2} \geqq 1$ and

$$\lambda^{-1}(\alpha) > \lambda_1^{-1}(\alpha) = \sup_{\phi \in H_{f_0}} \frac{(T_{f_0}\phi, \phi)_{f_0}}{(\phi, \phi)_{f_0}}. \tag{4.3}$$

By letting $f_0 = 1$ and $\phi = x^{1/4}$ and by performing the integration in (4.3), we obtain

$$\lambda^{-1}(\alpha) > \frac{16}{15} + \frac{16\alpha}{49}. \tag{4.4}$$

From (4.4) it is evident that as $\alpha \to +\infty$, $\lambda \to 0^+$. Since $b = \lambda^{2/3}$, we now have some estimate on the range of possible values of b for which the iteration is guaranteed to converge.

At each step of the iteration we have $\lambda_k(\alpha)$ is a continuous function of α for $\alpha \geqq -1$ (see [3]). If we utilize

$$\lambda_{k+1}^{-1} = \sup_{\phi \in H_{fk}} \frac{(T_{fk}\phi, \phi)_{fk}}{(\phi, \phi)_{fk}} \tag{4.5}$$

it is straightforward to show that $\lambda_k(\alpha)$ converges uniformly to $\lambda(\alpha)$ on compact subsets of $[-1, \infty)$ and thus $\lambda(\alpha)$ is a continuous function of α.

We conclude with some comments on the possible implementation of the iteration for actual computational purposes. This problem seems to be well adapted to an application of Galerkin's method. In Galerkin's method if $\{\phi_n\}$ is a complete orthonormal system in a separable Hilbert space H and if T is a compact selfadjoint linear operator on H with simple spectrum then for L_N the linear hull of $\{\phi_1, \cdots, \phi_N\}$ and P_N the orthogonal projection onto L_N, the eigenvalues and eigenfunctions of $P_N T$ converge to the eigenvalues and eigenfunctions of T as $N \to \infty$ (see [12]). If Galerkin's method is applied to the operator T_f of (2.4), then it is necessary after every iteration to reorthonormalize the sequence $\{\phi_n\}$ since the inner product in H_f changes with the generation of each new function f_k. It appears more appropriate to apply Galerkin's method to the operator M_f of (2.5) on $L^2[0, 1]$ and then use equation (2.6) to relate these eigenfunctions to the corresponding eigenfunctions of T_f. We are proceeding with actual numerical computations along this line using a complete orthonormal system of polynomials in $L^2[0, 1]$. The results of these investigations will be reported in a future paper.

REFERENCES

[1] V. RUSH AND S. H. CALDWELL, *Thomas-Fermi equation solution by the differential analyzer*, Phys. Rev., 38 (1931), no. 2, pp. 1898–1901.

[2] E. CODDINGTON AND N. LEVINSON, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[3] R. COURANT AND D. HILBERT, *Methods of Mathematical Physics*, vol. 1, Interscience, New York, 1953.

[4] P. CSAVINSKY, *Calculation of diamagnetic susceptibilities of ions using a universal approximate analytical solution of the Thomas-Fermi equation*, Bull. Amer. Phys. Soc., 2 18 (1973), no. 2, pp. 726–727.

[5] ———, *Universal approximate analytical solution of the Thomas-Fermi equation for ions*, Phys. Rev. A, 8 (1973), no. 3, pp. 1688–1701.

[6] H. T. DAVIS, *Introduction to Nonlinear Differential and Integral Equations*, Dover, New York, 1962.

[7] E. FERMI, *Un metodo statistico par la determinzione di alcune Proprietá dell' atome*, Rend. Accad. Naz. del Lincei, Cl. Sci. Fis., Mat. e. Nat., 6 (1927), no. 6, pp. 602–607.

[8] P. GOMBÁS, *Die statistische Theorie des Atoms*, Springer, Berlin, 1949.

[9] E. HILLE, *Some aspects of the Thomas-Fermi equation*, J. Analyse Math., 23 (1970), pp. 147–170.

[10] ———, *On a class of nonlinear second order differential equations*, Rev. Un. Mat. Argentina, 25 (1971), pp. 319–334.

[11] S. KARLIN, *Positive operators*, J. Math. Mech., 8 (1959), no. 6, pp. 907–937.

[12] M. KRASNOSELSKII, *Topological Methods in the Theory of Nonlinear Integral Equations*, Pergamon Press, New York, 1964.

[13] E. H. LIEB AND B. SIMON, *Thomas-Fermi theory revisited*, Phys. Rev. Lett., 31 (1973), pp. 681–683.

[14] ———, *The Thomas-Fermi theory of atoms, molecules, and solids*, Advances in Math., to appear.

[15] A. MAMBRIANI, *Su un teorema relativo alle equazioni differenziali ordinarie del 2^d ordine*, Rend. Accad. Naz. del Lincei, Cl. Sci. Fis., Mat. e Nat., 9 (1929), no. 6, 620–622.

[16] N. H. MARCH, *The Thomas-Fermi approximation in quantum mechanics*, Advances in Phys., 6 (1957), pp. 1–101.

[17] R. A. MOORE AND Z. NEHARI, *Nonoscillation theorems for a class of nonlinear differential equations*, Trans. Amer. Math. Soc., 93 (1959), pp. 30–52.

[18] R. V. RAMNATH, *A new analytical approximation for the Thomas-Fermi model in atomic physics*, J. Math. Anal. Appl., 31 (1970), pp. 285–296.

[19] F. RIESZ AND B. SZ-NAGY, *Functional Analysis*, Frederick Ungar, New York, 1955.

[20] A. SOMMERFELD, *Asymptotische Integration der Differential-Gleichung des Thomas-Fermischen Atoms*, Z. Phys., 78 (1932), pp. 283–308.

[21] Y. BAE SUH, *Perturbation calculation using Thomas-Fermi model*, Phys. Lett. A, 49 (1974), pp. 99–100.

[22] L. H. THOMAS, *The calculation of atomic fields*, Proc. Cambridge Philos. Soc., 23 (1927), pp. 542–548.

[23] J. S. WONG, *On the generalized Emden-Fowler equation*, SIAM Rev., 17 (1975), pp. 339–360.

SIAM J. MATH. ANAL.
Vol. 9, No. 3, June 1978

Copyright © 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0903-0008$1.00/0

SHORT PROOFS OF THREE THEOREMS ON ELLIPTIC INTEGRALS*

B. C. CARLSON†

Abstract. Duplication, reduction, and addition theorems are proved for a symmetric elliptic integral of the first kind by change of integration variable. Some comments are made on the reduction theorem and its relation to the addition theorem.

1. Half a dozen canonical forms for the elliptic integral of the first kind have been proposed by various authors [8, pp. 49–50], most recently [1]

$$R_F(x, y, z) = \frac{1}{2}\int_0^\infty [(t+x)(t+y)(t+z)]^{-1/2}\,dt. \tag{1}$$

By standardizing the interval of integration instead of the branch points of the integrand, one retains permutation symmetry in the finite branch points, thereby eliminating the linear transformations that plague Legendre's integral [4]. Such permutation symmetry is a property also of Weierstrass' canonical form, which is closely related to R_F [1, (3.10)]. The Landen transformation of (1) is discussed in references [1] and [2], the duplication theorem in [4] and [10], upper and lower approximations in [5], numerical properties in [2] and [9], and the addition theorem in [10, § 8]. Most of these may be compared with corresponding properties of Legendre's integral proved in [7], for instance.

A significant merit of (1) is the recently discovered reduction theorem,

$$\int_0^\infty [(t+a^2)(t+b^2)(t+c^2)(t+d^2)]^{-1/2}\,dt = \int_0^\infty [(t+x^2)(t+y^2)(t+z^2)]^{-1/2}\,dt, \tag{2}$$
$$x = ab+cd, \quad y = ac+bd, \quad z = ad+bc, \qquad a, b, c, d > 0,$$

which has important consequences for tables of elliptic integrals [6]. This formula was found by using the addition theorem, but its simplicity calls for a more direct proof. In the present note we shall prove (2) by a change of integration variable which allows other limits of integration without complicating the proof. The appropriate change of variable was found by a method which suggested similar proofs, also presented below, for the duplication and addition theorems. In each case permutation symmetry simplifies the algebraic manipulations. After proving the three theorems we shall say how the appropriate substitutions were found and comment further on the reduction theorem.

THEOREM 1 (Duplication theorem). *Let p, x, y, z be real numbers. Assume that $p+x$, $p+y$, $p+z$ are nonnegative and at most one of them is 0. Define*

$$q = p+(p+x)^{1/2}(p+y)^{1/2}+(p+x)^{1/2}(p+z)^{1/2}+(p+y)^{1/2}(p+z)^{1/2}, \tag{3}$$

all square roots being nonnegative. Then

$$\int_p^\infty [(s+x)(s+y)(s+z)]^{-1/2}\,ds = 2\int_q^\infty [(t+x)(t+y)(t+z)]^{-1/2}\,dt. \tag{4}$$

Proof. The appropriate change of integration variable is

$$t = [(s+x)^{1/2}+(s+y)^{1/2}][(s+x)^{1/2}+(s+z)^{1/2}]-x. \tag{5}$$

---

* Received by the editors July 21, 1976.

† Ames Laboratory–ERDA and Departments of Mathematics and Physics, Iowa State University, Ames, Iowa 50011.

Expansion of the product shows that the right side of (5) is symmetric in x, y, z and that $s=p$ implies $t=q$. For brevity define

$$X=(s+x)^{1/2}, \quad Y=(s+y)^{1/2}, \quad Z=(s+z)^{1/2}.$$

By symmetry,

$$t=(X+Y)(X+Z)-x=(Y+Z)(Y+X)-y=(Z+X)(Z+Y)-z,$$
$$(t+x)(t+y)(t+z)=(X+Y)^2(X+Z)^2(Y+Z)^2.$$

Differentiation of (5) yields

$$\frac{dt}{ds}=\tfrac{1}{2}(X^{-1}+Y^{-1})(X+Z)+\tfrac{1}{2}(X+Y)(X^{-1}+Z^{-1})=(2XYZ)^{-1}(X+Y)(X+Z)(Y+Z),$$

whence

$$[(s+x)(s+y)(s+z)]^{-1/2}\,ds=2[(t+x)(t+y)(t+z)]^{-1/2}\,dt.$$

Integration of this equation proves (4).

THEOREM 2 (Reduction theorem). *Let a, b, c, d, p be real nonnegative numbers and assume that at most one of a, b, c, d is 0. Define $x=ab+cd$, $y=ac+bd$, $z=ad+bc$, and*

(6) $$q=2[(p+a^2)(p+b^2)(p+c^2)(p+d^2)]^{1/2}+2p^2+p(a^2+b^2+c^2+d^2)-2abcd.$$

Then

(7) $$\int_0^p[(s+a^2)(s+b^2)(s+c^2)(s+d^2)]^{-1/2}\,ds=\int_0^q[(t+x^2)(t+y^2)(t+z^2)]^{-1/2}\,dt.$$

Proof. Let

(8) $$t=[(s+a^2)^{1/2}(s+b^2)^{1/2}+(s+c^2)^{1/2}(s+d^2)^{1/2}]^2-(ab+cd)^2.$$

Expansion of the squares shows that the right side of (8) is symmetric in a, b, c, d and that $s=p$ implies $t=q$. Define

$$A=(s+a^2)^{1/2}, \quad B=(s+b^2)^{1/2}, \quad C=(s+c^2)^{1/2}, \quad D=(s+d^2)^{1/2}.$$

By symmetry,

$$t=(AB+CD)^2-x^2=(AC+BD)^2-y^2=(AD+BC)^2-z^2,$$
$$(t+x^2)(t+y^2)(t+z^2)=(AB+CD)^2(AC+BD)^2(AD+BC)^2.$$

Differentiation of (8) yields

$$\frac{dt}{ds}=(AB+CD)(A^{-1}B+AB^{-1}+C^{-1}D+CD^{-1})$$
$$=(ABCD)^{-1}(AB+CD)(AC+BD)(AD+BC),$$

whence

$$[(s+a^2)(s+b^2)(s+c^2)(s+d^2)]^{-1/2}\,ds=[(t+x^2)(t+y^2)(t+z^2)]^{-1/2}\,dt.$$

THEOREM 3 (Addition theorem). *Let x, y, z be real nonnegative numbers and assume that at most one of them is 0. Let p be positive and define*

(9) $$q=p^{-2}[(p+x)^{1/2}(p+y)^{1/2}(p+z)^{1/2}+(xyz)^{1/2}]^2-p-x-y-z.$$

Then

$$(10)\qquad \int_p^\infty [(s+x)(s+y)(s+z)]^{-1/2}\,ds = \int_0^q [(t+x)(t+y)(t+z)]^{-1/2}\,dt.$$

Proof. Let

$$(11)\qquad t = xyz[s^{-1/2}(s^{-1}+x^{-1})^{1/2} + (s^{-1}+y^{-1})^{1/2}(s^{-1}+z^{-1})^{1/2}]^2 - x.$$

Expansion of the square shows that the right side of (11) is symmetric in x, y, z, and similar expansion of (9) shows that $s = p$ implies $t = q$. Define

$$W = s^{-1/2},\quad X = (s^{-1}+x^{-1})^{1/2},\quad Y = (s^{-1}+y^{-1})^{1/2},\quad Z = (s^{-1}+z^{-1})^{1/2}.$$

By symmetry,

$$t = xyz(WX+YZ)^2 - x = xyz(WY+ZX)^2 - y = xyz(WZ+XY)^2 - z,$$
$$(t+x)(t+y)(t+z) = (xyz)^3(WX+YZ)^2(WY+ZX)^2(WZ+XY)^2.$$

Differentiation of (11) yields

$$-s^2\frac{dt}{ds} = \frac{dt}{d(s^{-1})} = xyz(WX+YZ)(W^{-1}X + WX^{-1} + Y^{-1}Z + YZ^{-1})$$
$$= xyz(WXYZ)^{-1}(WX+YZ)(WY+ZX)(WZ+XY).$$

Since

$$WXYZ = s^{-2}(xyz)^{-1/2}[(s+x)(s+y)(s+z)]^{1/2},$$

it follows that

$$[(s+x)(s+y)(s+z)]^{-1/2}\,ds = -[(t+x)(t+y)(t+z)]^{-1/2}\,dt.$$

Integration of this equation completes the proof of (10).

2. All but one of the integrals in Theorems 1, 2, and 3 can be expressed in terms of the standard integral (1), since

$$(12)\qquad \int_p^\infty [(s+x)(s+y)(s+z)]^{-1/2}\,ds = 2R_F(x+p, y+p, z+p),$$

as one sees by taking $s-p$ as a new integration variable. Other integrals are reduced to this form by putting $\int_0^q = \int_0^\infty - \int_q^\infty$. Thus (10) can be rewritten in the form

$$(13)\qquad R_F(x+p, y+p, z+p) + R_F(x+q, y+q, z+q) = R_F(x, y, z).$$

The sum of two elliptic integrals on the left side is the sum of two arguments in the addition theorem for elliptic functions. The symmetry of (13) in p and q matches the symmetry of the rationalized form of (9),

$$(14)\qquad (pq - xy - xz - yz)^2 = 4xyz(p+q+x+y+z).$$

The special case $p = q$ of (13) and (14) coincides with the special case $p = 0$ of (4),

$$(15)\quad R_F(x, y, z) = 2R_F(x+\lambda, y+\lambda, z+\lambda),\qquad \lambda = x^{1/2}y^{1/2} + x^{1/2}z^{1/2} + y^{1/2}z^{1/2}.$$

The rationalized form of (3) is

$$(16)\qquad 2(2p+q+x+y+z) = \frac{(q+y)(q+z)}{q+x} + \frac{(q+z)(q+x)}{q+y} + \frac{(q+x)(q+y)}{q+z}.$$

Substitution (5) was found by combining the previously known (15) with (12) to obtain

$$\int_p^\infty [(s+x)(s+y)(s+z)]^{-1/2}\,ds = 2R_F(x+p,\,y+p,\,z+p)$$
$$= 4R_F(x+p+\sigma,\,y+p+\sigma,\,z+p+\sigma)$$
$$= 2\int_q^\infty [(t+x)(t+y)(t+z)]^{-1/2}\,dt,$$

where $q = p+\sigma$ and the second equation of (15) gives

$$\sigma = (x+p)^{1/2}(y+p)^{1/2} + (x+p)^{1/2}(z+p)^{1/2} + (y+p)^{1/2}(z+p)^{1/2}.$$

We have arrived at (3), which is essentially the same as (5).

Substitution (8), with s and t replaced by p and q, was found by putting $s = t+p$ in

$$\int_p^\infty [(s+a^2)(s+b^2)(s+c^2)(s+d^2)]^{-1/2}\,ds,$$

applying the previously known (2), adding q to the variable of integration, and requiring p and q to vanish together.

In [6] the reduction theorem was proved by using the addition theorem. Conversely Theorem 3 can be proved directly from Theorem 2, and substitution (11) was found in the course of doing so, as follows. On the left side of (10) we take s^{-1} as a new variable of integration to get an integral having the same form as the left side of (7) with $a = 0$. The limit of integration on the right side of (7) is then found from (8) more conveniently than (6). Multiplying the variable of integration on the right side by xyz (in the notation of Theorem 3) yields both (10) and (11).

Although we have assumed all quantities to be real, Theorems 1, 2, and 3 remain valid for complex values by the permanence of functional relations provided no singularities of the integrands are encountered. In particular $R_F(x, y, z)$ as defined by (1) is holomorphic if x, y, z lie in the complex plane cut along the nonpositive real axis.

3. The reduction theorem has an interpretation in terms of capacities, since each side of (2) is twice the reciprocal capacity of an ellipsoid [3, (4.2)]. The theorem asserts that an ellipsoid in $\mathbb{R}^4$ with semiaxes a, b, c, d has the same capacity as an ellipsoid in $\mathbb{R}^3$ with semiaxes $ab+cd$, $ac+bd$, $ad+bc$.

One does not expect (2) to have an analogue in more than four variables. The group S_4 of permutations of a, b, c, d has an invariant subgroup V consisting of the identity and three products of two-cycles: $(ab)(cd)$, $(ac)(bd)$, and $(ad)(bc)$. The permutations belonging to V leave unchanged the combinations $x = ab+cd$, $y = ac+bd$, $z = ad+bc$, while the factor group S_4/V is isomorphic to the group S_3 of permutations of x, y, z. It is V which makes S_4 solvable and in Galois theory allows the general quartic equation to be solved by radicals. Since S_n is not solvable if $n > 4$, one cannot expect to find analogous combinations of more than four variables.[1]

Although the polynomials in the reduction theorem are explicit products of linear factors, one can readily deduce the following unfactored version involving quartic and cubic polynomials with the same invariants.

THEOREM 2A. *Let a be real and strictly positive. Assume that the quartic polynomial $as^4+bs^3+cs^2+ds+e$ is strictly positive if $s > 0$ and either is positive or has a simple*

[1] I thank Professor Irvin Hentzel for suggesting this argument.

zero at $s=0$. *Assume* $0 \leqq p \leqq \infty$ *and define*

$$f=c+6a^{1/2}e^{1/2}, \quad g=bd+4ca^{1/2}e^{1/2}+8ae, \quad h=(be^{1/2}+da^{1/2})^2, \tag{17}$$
$$q=2a^{1/2}(ap^4+bp^3+cp^2+dp+e)^{1/2}+2ap^2+bp-2a^{1/2}e^{1/2}.$$

Then

$$\int_0^p (as^4+bs^3+cs^2+ds+e)^{-1/2}\,ds=\int_0^q (t^3+ft^2+gt+h)^{-1/2}\,dt. \tag{18}$$

Proof. Put $s=a^{-1/2}\sigma$ so that the polynomials on both sides of (18) become monic. They can now be identified with the polynomials in (7), and coefficients can be compared.

REFERENCES

[1] B. C. CARLSON, *Normal elliptic integrals of the first and second kinds*, Duke Math. J., 31 (1964), pp. 405–419.
[2] ———, *On computing elliptic integrals and functions*, J. Math. and Phys., 44 (1965), pp. 36–51.
[3] ———, *Some inequalities for hypergeometric functions*, Proc. Amer. Math. Soc., 17 (1966), pp. 32–39.
[4] ———, *Hidden symmetries of special functions*, SIAM Rev., 12 (1970), pp. 332–345.
[5] ———, *Inequalities for a symmetric elliptic integral*, Proc. Amer. Math. Soc., 25 (1970), pp. 698–703.
[6] ———, *Elliptic integrals of the first kind*, this Journal, 8 (1977), pp. 231–242.
[7] A. CAYLEY, *Elliptic Functions*, 2nd ed., Dover, New York, 1961.
[8] F. KLEIN, *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, Springer-Verlag, Berlin, 1926.
[9] W. J. NELLIS AND B. C. CARLSON, *Reduction and evaluation of elliptic integrals*, Math. Comput., 20 (1966), pp. 223–231.
[10] D. G. ZILL AND B. C. CARLSON, *Symmetric elliptic integrals of the third kind*, Ibid., 24 (1970), pp. 199–214.

SIAM J. MATH. ANAL.
Vol. 9, No. 3, June 1978

Copyright © 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0903-0009$1.00/0

ASYMPTOTIC BEHAVIOR OF SOME DETERMINISTIC EPIDEMIC MODELS*

FRANK J. S. WANG†

Abstract. We consider the asymptotic behavior of the solution of a system of nonlinear Volterra integral equations which arises in study of the spread of a disease for which it is assumed that the rate of a susceptible becoming infected depends only on the proportion of infectives, the rate of an infected individual recovering from the disease depends only on the length of time since the disease was contracted, and that recovered individuals are permanently immune from further attack.

We examine the limiting behavior of the solution which gives us results on the total size of the epidemic, i.e., the proportion of the total number of individuals that finally contracts the disease.

1. Introduction. Deterministic models have a long history of use in the description of the spread of an infection. A basic reference is the monograph of Bailey [1] which contains a description of both stochastic and deterministic epidemic models. Many widely used models assume division of a fixed population into three disjoint classes: susceptibles, infectives and individuals who are removed from the susceptible infective interaction by isolation, death or permanent immunity due to previous infection. However, most of these models make the assumption that the infection rate is proportional to the proportion of infectives and the removal rate is independent of how long the individual has had the disease. These assumptions are clearly far from being realistic. Wang [9], [10] suggests a general model for the study of the spread of disease and examines the relationship between the stochastic and deterministic versions. The present paper studies the asymptotic behavior of the deterministic model proposed by Wang.

Let us suppose that a population of size N has at any time t a proportion $s(t)$ of individuals susceptible to a certain disease, a proportion $x(t)$ of individuals actually infected, and a proportion $y(t)=1-s(t)-x(t)$ of recovered individuals that are permanently immune from further attack. We assume that the probability of a particular susceptible individual becoming infected during $(t, t+dt)$ is $\alpha(x(t))\,dt$ for some function α such that $\alpha(0)=0$, and that the probability of an infected individual staying infected for at least a length of time t is $F(t)$ for some decreasing function $F(t)$ such that $F(0)=1$. Wang [9] constructs a process with the above properties and proposes a continuous deterministic model given by the system of equations

$$
\begin{aligned}
x(t)&=r(t)+\int_0^t \alpha(x(u))s(u)F(t-u)\,du,\\
s(t)&=1-x(0)-\int_0^t \alpha(x(u))s(u)\,du.
\end{aligned}
\tag{1}
$$

Here, $r(t)$ represents the proportion of infectives who are already infected at time zero and are still infective at time t. This model generalizes Bailey's general epidemic model (Bartlett [2]) in which he makes the traditional assumption that the infection rate is proportional to the proportion of infectives, i.e., $\alpha(x)=\alpha\cdot x$ for some constant α, and F is exponential with parameter β. In Bailey's general model, the system (1) reduces to

$$
\begin{aligned}
x'(t)&=\alpha\cdot x(t)\cdot s(t)-\beta x(t),\\
s'(t)&=-\alpha x(t)s(t).
\end{aligned}
$$

* Received by the editors April 23, 1976, and in revised form November 10, 1976.
† Department of Mathematics, University of Montana, Missoula, Montana 59801.

This system has been studied by many authors (e.g. Kermack and McKendrick [7], Bailey [1], Kendell [6], Waltman [8], Hoppensteadt [5], Hethcote [4]). In studying the behavior of solution of (1), we are interested in the asymptotic behavior of $y(t)$.Denote the limit of $y(t)$ as $t \to \infty$ by y provided it exists. Then y is the proportion of the total number of individuals that finally contracts the disease. It is usually called the size or the intensity of the epidemic (see Waltman [8]). We give conditions under which y is strictly less than one and obtain an upper (lower) bound for y by approximating α from above (below) by a convex (concave) function $\alpha^+(\alpha^-)$. When $\alpha(x) = \alpha \cdot x$ is linear, the total size of the epidemic is the unique positive solution of the equation

$$y + c \cdot e^{-my} - 1 = 0,$$

where

$$m = \alpha \int_0^\infty F(s)\, ds$$

is the inverse of the generalized relative removal rate and

$$c = (1 - x(0)) \exp\left\{ mx(0) - \alpha \cdot \int_0^\infty r(t)\, dt \right\}.$$

In this case, our model is a special case of a more general model (SIER) studied by F. Hoppensteadt [5] and P. Waltman [8]. In the SIER model, the class S is modified to be "susceptible and unexposed to the infection", and a class (E) of "exposed but not yet infective" individuals is added to the three disjoint classes x, s and y. The limiting behavior of the functions $s(t)$, $x(t)$, $E(t)$ and $y(t) = 1 - s(t) - x(t) - E(t)$ of the SIER model are considered in the monograph by P. Waltman [8].

2. Limits of solution. We shall make the following assumptions on the functions α, r and F:

A_1: We assume that $\alpha(0) = 0$, that $\alpha(x)$ is continuous and nonnegative on $[0, 1]$, and that $\alpha'(x)$ is continuous on $[0, 1]$.

A_2: We assume that $r(t)$ is a nonincreasing continuous function, and that $r(t) \to 0$ as $t \to \infty$. We also assume that $0 \leqq r(t) \leqq 1$ and $\int_0^\infty r(t)\, dt < \infty$.

A_3: We assume that $F(0) = 1$, that $F(t)$ is nonnegative monotonic nonincreasing and tends to zero as $t \to \infty$. We also assume that $\int_0^\infty F(s)\, ds < \infty$.

THEOREM 1. *Suppose the conditions* A_1, A_2 *and* A_3 *are satisfied. Then the system* (1) *has a unique solution pair* $(x(t), s(t))$, *where* $x(t)$ *and* $s(t)$ *are nonnegative, bounded by* 1. *Furthermore,* $x(t)$ *tends to zero as* $t \to \infty$ *and* $s(t)$, *being a monotonic decreasing function, tends to some limit* s *as* $t \to \infty$.

Proof. Let us first assume that α is Lipschitz continuous on $[0, \infty)$. Then the existence and the uniqueness of the solution can be proved by a standard contraction method (e.g. Wang [9, Lemma 5.2]). Since it follows from (1) that

$$s(t) = (1 - x(0)) \exp\left\{ -\int_0^t \alpha(x(u))\, du \right\}$$

and $1 - x(t) \geqq s(t)$, it is clear that $s(t)$ and $x(t)$ are nonnegative, bounded by 1, and that $\lim_{t\to\infty} s(t) = s$ exists. Suppose $s = 0$. Then A_2 and A_3 imply that $x(t) \to 0$ as $t \to \infty$. Suppose $s > 0$. Then

$$\int_0^\infty \alpha(x(t))\, dt < \infty.$$

This, together with (1) and A_3, again implies $x(t) \to 0$ as $t \to \infty$.

Since $0 \leqq x(t) \leqq 1$, the solution of (1) is independent of the behavior of $\alpha(x)$ over the set of all $x > 1$. Thus, by redefining α on $(1, \infty)$ if necessary, we could assume that α is Lipschitz continuous on $[0, \infty)$. This completes the proof.

One of the classical results concerning the asymptotic behavior of the solution of the renewal equation

$$z(t) = g(t) + \int_0^t z(s) f(t-s)\, ds \tag{2}$$

where f and $g \not\equiv 0$ are given nonnegative functions, is that $z(t)$ grows exponentially as $t \to \infty$ if $g(t) \to 0$, f is of bounded total variation over $[0, \infty)$ and

$$\int_0^\infty f(t)\, dt > 1$$

(Feller [3]). The proof of the next theorem uses this well-known result.

THEOREM 2. *Suppose the hypotheses* A_1, A_2 and A_3 *are satisfied. Let* $(x(t), s(t))$ *be the unique solution pair of the system* (1). *Then* $s(t)$ *tends to a nonzero limit* s *as* $t \to \infty$, *and*

$$0 < s \leqq \left(\alpha'(0) \cdot \int_0^\infty F(u)\, du\right)^{-1}.$$

Proof. Choose a number $M > 0$ such that $\alpha'(0) > M$, and define $\phi(x) = \alpha(x) - Mx$. Since $\phi(0) = 0$ and $\phi'(0) > 0$, $\phi(x)$ is positive for all sufficiently small $x \geqq 0$. Since $x(t) \to 0$ as $t \to \infty$, there exists a number $T > 0$ such that $\phi(x(t)) \geqq 0$ for all $t \geqq T$.

We now rewrite the first equation in (1) as

$$x(t) = h(t) + \int_T^t M \cdot s \cdot x(u) F(t-u)\, du \tag{3}$$

where

$$h(t) = r(t) + \int_0^T \alpha(x(u)) s(u) F(t-u)\, du$$
$$+ \int_T^t [\alpha(x(u)) \cdot (s(u) - s) + s \cdot \phi(x(u))] \cdot F(t-u)\, du.$$

Make a change of variable $t - T = c$ and define $z(c) = x(c+T)$, $g(c) = h(c+T)$ for $c \geqq 0$; then the equation (3) for $t \geqq T$ becomes

$$z(c) = g(c) + \int_0^c M \cdot s \cdot z(u) F(c-u)\, du.$$

Since $s(t)$ decreases to s monotonically and $\phi(x(t)) \to 0$ as $t \to \infty$, $h(t)$ is nonnegative and tends to zero as $t \to \infty$. This implies g is nonnegative and $g(t) \to 0$ as $t \to \infty$. Since $z(t) = x(t+T) \to 0$ as $t \to \infty$, the well-known classical result on the linear renewal equation (2) implies that

$$\int_0^\infty Ms \cdot F(t)\, dt \leqq 1, \quad \text{or} \quad s \leqq \left(M \int_0^\infty F(t)\, \mathrm{d}t\right)^{-1}.$$

Since the above inequality is true for all M such that $\alpha'(0) > M$, the same inequality holds with M replaced by $\alpha'(0)$. To show that $s > 0$, we choose a number $N > 0$ such that $\alpha'(0) < N$ and define $\Psi(x) = \alpha(x) - N \cdot x$. Since $\Psi(0) = 0$ and $\Psi'(0) < 0$, $\Psi(x) < 0$ for all sufficiently small $x \geqq 0$. This implies

$$\int_T^\infty \Psi(x(t))\,dt<0 \tag{4}$$

for T large enough. Integrating both sides of the first equation in (1), we see that it follows from the Fubini theorem that

$$\int_0^\infty x(t)\,dt=\int_0^\infty r(t)\,dt+(1-x(0)-s)\left(\int_0^\infty F(t)\,dt\right)<\infty.$$

Thus

$$0<\int_T^\infty x(t)\,dt<\infty \tag{5}$$

for all $T \geqq 0$. Since

$$0\leqq\int_T^0 \alpha(x(t))\,dt=N\cdot\int_T^\infty x(t)\,dt+\int_T^\infty \psi(x(t))\,dt,$$

inequalities (4) and (5) imply that the integral on the left side of the preceding equality is finite. This implies

$$\int_0^\infty \alpha(x(t))\,dt<\infty$$

and hence

$$s=(1-x(0))\cdot\exp\left\{-\int_0^\infty \alpha(x(t))\,dt\right\}>0.$$

This establishes the theorem.

3. The total size of the epidemic. In this section we examine the size of the epidemic $y=1-s$. It is clear from Theorem 2 that

$$y\geqq 1-\left(\alpha'(0)\cdot\int_0^\infty F(t)\,dt\right)^{-1}.$$

A different lower bound and an upper bound for y can be obtained. We define

$$k(t)=\exp\left\{-\int_0^t \alpha(x(u))\,du\right\};$$

then

$$s(t)=(1-x(0))k(t) \quad \text{and} \quad (1-x(0))k(\infty)=1-y.$$

Let α^+ be an arbitrary convex function such that

$$\alpha^+(x)\geqq\alpha(x) \quad \text{for all } x\in[0,1],$$

and define

$$m=\int_0^\infty \alpha^+(F(u))\,du,$$

$$a=x(0)\int_0^\infty \alpha^+(r(t)/x(0))\,dt.$$

Then it follows from the convexity of α^+ and Jensen's inequality that

$$\alpha(x(t)) \leqq \alpha^+(x(t)) \leqq x(0)\alpha^+(r(t)/x(0)) + (1-x(0)) \int_0^t \alpha^+(F(t-u))\, d[1-k(u)].$$

Integrating both sides of the above inequality and applying the Fubini theorem, we obtain

$$\int_0^\infty \alpha(x(t))\, dt \leqq a + m(1-x(0)) \cdot (1-k(\infty)).$$

This implies

$$0 \leqq y + c\, e^{-my} - 1$$

where

$$c = (1-x(0)) \cdot \exp\{mx(0) - a\}.$$

Let

$$p(y) = y + c\, e^{-my} - 1.$$

Then $p(1) > 0$, $p(x(0)) < 0$ and $p'' > 0$. By elementary steps we conclude that $p(y) = 0$ has a unique solution y^+ such that $1 \geqq y^+ \geqq x(0)$ and that $y \leqq y^+$. Let α^- be a concave function such that $\alpha^- \leqq \alpha$ for all $x \in [0, 1]$. An argument similar to the preceding one shows that $y^- \leqq y$ where y^- is the unique root in $[x(0), 1]$ of the equation

$$y + \bar{c}\, e^{-\bar{m}y} - 1 = 0,$$

with

$$\bar{m} = \int_0^\infty \alpha^-(F(u))\, du,$$

and

$$\bar{c} = (1-x(0)) \exp\left\{x(0) \cdot \left[\bar{m} - \int_0^\infty \alpha^-(r(t)/x(0))\, dt\right]\right\}.$$

In particular if $\alpha(x) = \alpha \cdot x$ is linear, y is the solution of the equation

$$y + c \cdot e^{-my} - 1 = 0$$

where m and c are as stated at the end of the introduction.

Acknowledgment. The author would like to express his thanks to the referee, for pointing out the results of P. Waltman and F. Hoppensteadt.

REFERENCES

[1] NORMAN T. J. BAILEY, *The Mathematic Theory of Epidemic*, Hafner, New York, 1957.

[2] M. S. BARTLETT, *Stochastic Population Models in Ecology and Epidemiology*, Methuen, London, 1960.

[3] W. FELLER, *On the integral equation of renewal theory*, Ann. Math. Statist., 12 (1941), pp. 243–267.

[4] HERBERT HETHCOTE, *Qualitative analysis of communicable disease models*, Math. Biosci., 28 (1976), pp. 335–356.

[5] FRANK HOPPENSTEADT, *Mathematical theories of population: Demographics, Genetics and Epidemics*, SIAM Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, 1975.

[6] D. G. KENDALL, *Deterministic and Stochastic epidemic in closed population*, Proc. Third Berkeley Symposium on Math. Stat. and Prob., vol. 4, University of California Press, Berkeley, 1956, pp. 149–165.

[7] W. O. KERMACK AND A. G. MCKENDRICK, *A contribution to the mathematical theory of epidemics*, Proc. Roy. Soc. Ser. A., 115 (1927), pp. 700–721.

[8] PAUL WALTMAN, *Deterministic threshold models in the theory of epidemics*, Lecture Notes in Biomathematics, Vol. 1, Springer-Verlag, New York, 1974.

[9] FRANK J. S. WANG, *Limit theorems for age and density dependent stochastic population models*, J. Math. Biol., 2 (1975), pp. 373–400.

[10] ———, *Gaussian approximation of some closed stochastic epidemic models*, J. Appl. Probability, 14 (1977), pp. 221–231.

SIAM J. MATH. ANAL.
Vol. 9, No. 3, June 1978
Copyright © 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0903-0010$1.00/0

ON THE INTERVAL OF EXISTENCE FOR NONLINEAR TWO POINT BOUNDARY VALUE PROBLEMS*

PAUL B. BAILEY† AND MICHAEL J. NORRIS†

Abstract. The existence question for a nonlinear, two point boundary value problem (continuous and Lipschitzian) can be reduced to that of the corresponding unforced problem with one zero boundary condition. For fixed left-hand end point, the set of right-hand end points for which existence fails consists of isolated points and isolated intervals.

Certain bounding curves arise in the case of failure of existence for an interval of right-hand end points, and smoothness and dependence on initial conditions of these curves are determined in part. Also some relationships between the given problem and the problem with initial and terminal points reversed are established.

1. Introduction. In the case of a linear two point boundary value problem

$$y''+py'+qy=\varphi \tag{1}$$

$$y(a)=A, \qquad y(b)=B \tag{2}$$

with p, q, and φ continuous on $[a, b]$, the question of whether or not there exists a solution for arbitrary real numbers A and B and arbitrary real valued functions can be answered in terms of the so-called homogeneous problem

$$y''+py'+qy=0 \tag{3}$$

$$y(a)=0, \qquad y(b)=0. \tag{4}$$

Thus, if the homogeneous problem (3), (4) has only the trivial solution $y(x)\equiv 0$, then the inhomogeneous problem (1), (2) has one, and only one, solution for arbitrary A, B, and $\varphi(x)$.

The same question for nonlinear equations of the form

$$y''(x)+f(x, y(x), y'(x))=0, \tag{5}$$

where f is continuous and satisfies a Lipschitz condition in y and y'—i.e., there exist constants K and L such that

$$|f(x, y, y')-f(x, z, z')|\leqq K|y-z|+L|y'-z'| \tag{6}$$

for all x, y, y', z, z'—is not so simply answered, of course. For one thing, uniqueness is not so intimately associated with existence in the nonlinear case as it is in the linear. Nevertheless the existence question for the problem (5), (2) can be answered in terms of the corresponding "unforced" (we can no longer say "homogeneous") problem

$$y''(x)+f(x, y(x), y'(x))-f(x, 0, 0)=0 \tag{7}$$

$$y(a)=0, \qquad y(b)=B. \tag{8}$$

It will be shown that (5), (2) has at least one solution for every A and B if and only if (7), (8) has at least one solution for every B. Furthermore, the set of points b for which this holds consists of isolated points and isolated intervals. Specifically, given b there exists a positive δ such that existence fails throughout $(b, b+\delta)$ or existence holds

* Received by the editors October 29, 1974, and in final revised form November 29, 1976.

† Applied Mathematics Department, Sandia Laboratories, Albuquerque, New Mexico 87115. This work was supported by the United States Energy Research and Development Administration (ERDA) under Contract AT(29-1)-789.

throughout $(b, b+\delta)$ and, at the same time, existence fails throughout $(b-\delta, b)$ or existence holds throughout $(b-\delta, b)$.

Existence failure at an isolated right-hand end point b is, of course, the only type of failure possible for linear problems. However, since failure for a whole interval of end points can easily happen in nonlinear problems, this situation is examined in more detail, some of the main results being illustrated by several examples. Certain bounding curves arise in the case of failure of existence on an interval, and smoothness characteristics and dependence on initial conditions of these curves are determined in part. Some relationships between the given problem and the problem with initial and terminal points reversed are established, primarily for the symmetric case corresponding to the first boundary value problem (FBVP) with conditions of form (2).

Although we discuss only the first boundary value problem, all results, except Theorem 3(iii) are also valid for the second boundary value problem (SBVP) where the boundary conditions are of the form

$$y'(a)=m, \qquad y(b)=B. \tag{2'}$$

At most trivial changes are needed in the statements or proofs of the theorems.

2. Preliminaries. It will be assumed throughout that the function $f(x, y, y')$ appearing in (5) is real valued, continuous, and satisfies a Lipschitz condition (6) on any finite interval relevant to the discussion. This means, in particular, that the class of functions being considered includes linear functions as in (1). For this class of Lipschitz functions it is well known [1] that initial value problems (IVP) for (5) always have unique solutions, that these solutions exist on the whole real line, and that they and their first derivatives on any finite interval depend continuously, uniformly on any interval, on their initial conditions.

Another important and well-known [2] result which will be needed several times is that any first or second boundary value problem for (5) (when f is continuous and Lipschitzian) has one, and only one, solution if the interval $[a, b]$ is sufficiently small; how small depends only upon the Lipschitz constants. One obvious consequence of this is the fact that the solution $y(x; m)$ of (5) and $y(a)=A$, $y'(a)=m$ has $y(b; m)\to \pm\infty$ and $y'(b; m)\to\pm\infty$ as $m\to\pm\infty$ for $b-a$ positive and sufficiently small and $y(b; m)\to\mp\infty$, $y'(b; m)\to\pm\infty$ for $b-a$ negative and sufficiently small.

Moreover, in terms of the norm defined by

$$\|u\|(x)=|u(x)|+|u'(x)| \tag{9}$$

one has the following useful result:

LEMMA 1. *If $\varphi(x)$, $\psi(x)$ are continuous, and if $u(x)$, $v(x)$ satisfy*

$$u''(x)+f(x, u(x), u'(x))=\varphi(x) \tag{10}$$

$$v''(x)+f(x, v(x), v'(x))=\psi(x) \tag{11}$$

on an interval $[a, b]$, then

$$\|u-v\|(x)\leqq\|u-v\|(a)\cdot\exp(k|x-a|)+k^{-1}(M+N)\cdot(\exp(k|x-a|)-1), \tag{12}$$

$$k=\max\{K, L+1\}, \quad M=\max_{x\in[a,b]}|\varphi(x)|, \quad N=\max_{x\in[a,b]}|\psi(x)|.$$

Proof. In the terminology of [1, pp. 8, 19], u and v are ε-approximate solutions of (5) with $\varepsilon=M$ and N, respectively, □

3. Existence intervals. Since solutions to initial value problems for (5) are continuous functions of the initial conditions, the set of values which can be obtained at $x=b$ by solutions of (5) which satisfy the fixed initial condition, $y(a)=A$, constitutes an interval of real numbers. This interval may be a finite interval (possibly a point), a half line, or the whole real line. In the last case we say that existence holds at $x=b$. (In the linear case the only possibilities are a single point or the whole real line.) The following theorem shows that the nature of this interval of values attainable at any particular point b is independent of the initial value, A, at $x=a$ and also independent of the "forcing term," $-f(x,0,0)$ (which in the linear case (1) was denoted by $\varphi(x)$).

THEOREM 1. *If for a particular A and $\varphi(x)$ there exists a number $\bar{A}$ (or $\underline{A}$) such that every solution, $y(x)$, of*

$$y''(x)+f(x, y(x), y'(x))=\varphi(x) \tag{13}$$

$$y(a)=A \tag{14}$$

also satisfies $y(b)\leqq\bar{A}$ ($\geqq\underline{A}$), then for any A and any φ there exists an $\bar{A}$ (or $\underline{A}$) for which the inequality holds.

Proof. Denote by $y(x; a, A, m, \varphi)$ the solution of (13), (14) and $y'(a)=m$. Suppose $y(b; a, A, m, \varphi)\leqq\bar{A}$ for all m. By Lemma 1,

$$|y(b; a, A', m, \psi)-y(b; a, A, m, \varphi)| \leqq|A-A'|\cdot E+k^{-1}(M+N)(E-1), \quad E=\exp(k|b-a|);$$

hence

$$\begin{aligned} y(b; a, A', m, \psi) &\leqq y(b; a, A, m, \varphi)+|A-A'|\cdot E+k^{-1}(M+N)(E-1)\\ &\leqq\bar{A}+|A-A'|\cdot E+k^{-1}(M+N)(E-1), \end{aligned}$$

where the right side does not depend on m. Obviously the same simple argument works in the case that $y(b; a, A, m, \varphi)\geqq\underline{A}$. □

When the set of values which can be reached at $x=b$ includes a half line containing all positive (negative) numbers of large magnitude we will say that upper existence (lower existence) holds at b.

Unlike the linear case, wherein existence fails only for a discrete set of right-hand end points, existence can fail for a whole interval of end points in the nonlinear case. An easy example is the case

$$y''+|y|=0, \quad y(0)=0, \quad y(b)=B. \tag{15}$$

It is easy to write down the general solution to this differential equation, and it is obvious that upper existence fails whenever $b\geqq\pi$. (See Fig. 1.1 of [2, p. 9]).

However, even nonlinear problems cannot have existence fail at a set of right-hand end points having a limit point, b, unless existence fails on a whole interval of points with end point b.

THEOREM 2. *If upper (lower) existence for* (5), (14) *fails at $x=b_1, b_2, \cdots, b_n, \cdots$ having a finite limit point, b_0, then there is an interval of points (b_0, b^*), or (b^*, b_0), such that upper (lower) existence fails at every point of the interval.*

Proof. Suppose the theorem is false. Suppose, for instance, that $b_n\to b_0$ from the right, that to each b_n there corresponds a $\bar{B}_n$ such that every solution $y(x)$ of (5), (14) passes below $\bar{B}_n$, i.e., $y(b_n)\leqq\bar{B}_n$, and that between every b_n, b_{n+1} there is some c_n for which upper existence holds.

Now we know that if n is large enough, b_n, b_{n+1} are sufficiently close together that any first boundary value problem on $[b_{n+1}, b_n]$ has a unique solution. In particular, there is a unique solution, $y_n(x)$, of (5) through the points $(b_{n+1}, \bar{B}_{n+1})$ and $(b_n, \bar{B}_n)$. Choose a number C_n such that the point (c_n, C_n) lies above $(c_n, y_n(c_n))$. Since upper existence is being assumed to hold at c_n, there is a solution $y(x)$ of (5), (14) satisfying $y(c_n)=C_n$. But since upper existence fails at b_{n+1} and b_n, this solution passes below $y_n(x)$ at b_{n+1} and b_n. Necessarily this $y(x)$ must meet $y_n(x)$ at least twice in the interval $[b_{n+1}, b_n]$, which violates the uniqueness for first boundary value problems that is known to hold if $[b_{n+1}, b_n]$ is small enough.

The contradiction establishes the theorem in this case, and obviously the other cases can be treated in exactly the same way. □

Theorem 2 is still valid if (14) is replaced by

$$y'(a)=m. \tag{14'}$$

However, Theorem 2 need not hold in the case of the second boundary value problem (5), (14), $y'(b)=n$. For, as the following example shows, one can have a linear equation of form (5) for which a can be chosen so that any solution, y, satisfying (14) (or (14')) has $y'(b_n)=0$ for a sequence of points $\{b_n\}$ converging monotonely to b_0 from above while for b in (b_{n+1}, b_n) as y runs through the indicated family of solutions $y'(b)$ runs through all real numbers.

Example 1. Let C be any function which is continuously differentiable on (c, ∞) with C' having a limit from the right at c. Then there exist continuous functions, f and g, on $(-\infty, \infty)$ and a number, a, such that for any solution on $(-\infty, \infty)$ of

$$y''=fy+gy'$$

$y'-Cy$ is constant on (c, ∞) and the derivative of that solution with $y(a)=0$, $y'(a)=1$ has zeros on (c, ∞) at exactly the zeros of C. Since the equation is linear, existence will fail on (c, ∞) at the zeros of C and hold elsewhere on (c, ∞).

Proof. C also has a limit from the right at c. Take an extension of C (leaving notation unchanged) which is continuously differentiable on an open interval (d, ∞), with $d<c$, and satisfies $C(b)=0$, $C'(b)<0$ at some point, b, of (d, c).

Let f be equal to C' on $[b, \infty)$ and $C'(b)$ on $(-\infty, b)$. Let g be equal to C on $[b, \infty)$ and $C(b)$, so 0, on $(-\infty, b)$. Then f and g are continuous on $(-\infty, \infty)$.

If y is a solution of the differential equation, then on $[b, \infty)$ one has $y''=C'y+Cy'$, $(y'-Cy)'=0$. For $A\neq 0$ let $y(x; A)$ be that solution with $y(c; A)=A$ and $y'(c; A)=C(c)A$. Then

$$y'(x; A)=C(x)y(x; A),$$

$$y(x; A)=A \exp \int_c^x C(\xi)\, d\xi.$$

Since $y(x; A)$ has no zeros on $[b, \infty)$, $y'(x; A)$ has the same zeros there as $C(x)$.

Now $y'(b; A)=C(b)y(b; A)=0$, so on $(-\infty, b)$

$$y(x; A)=y(b; A)\cos(\sqrt{-C'(b)}(x-b)).$$

Thus for $a=b-\pi(2\sqrt{-C'(b)})^{-1}$ one has $y(a; A)=0$ and $y'(a; A)=y(b; A)\sqrt{-C'(b)}=A \exp(-\int_b^c C(\xi)\, d\xi)\sqrt{-C'(b)}$. Taking A so that $y'(a; A)=1$ yields $y(x; A)$ as the solution which is 0 and has derivative 1 at a. □

4. Bounding curves.

DEFINITION. If upper (lower) existence fails on a proper interval I, then corresponding to each initial value, A, at the fixed end point $x=a$ there is an *upper* *(lower) bounding curve*, $y=\bar{B}(x)$ $(\underline{B}(x))$, finite, defined on I by

$$\bar{B}(x)=\sup\{y(x)\colon y \text{ is a solution of (5), (14)}\}$$

$$(\underline{B}(x)=\inf\{y(x)\colon y \text{ is a solution of (5), (14)}\}).$$

Note that while upper and lower existence can both fail at an isolated point, they cannot both fail over the same proper interval, as the following theorem shows.

THEOREM 3. *If all solutions of* (5) *through some point,* (a, A), *are bounded both above and below at* $x=b$, *then*

(i) *all solutions through* (b, B) *are bounded above and below at* $x=a$,

(ii) b *is an isolated point of failure of existence,*

(iii) a *is an isolated point of failure of existence for* b *taken as the initial point.*

Proof. Suppose all solutions through (a, A) are bounded above by $\bar{B}$ and below by $\underline{B}$ at $x=b$. Let B be fixed with $B>\bar{B}$. Since no solution through (b, B) can go through (a, A), and since two such solutions with one going above (a, A) and one going below (a, A) would lead to such a solution through (a, A), then either A is an upper bound or A is a lower bound for all such solutions at $x=a$. We may assume without loss of generality in what follows that A is an upper bound.

Consider the mapping T which takes initial conditions, (A', m'), at $x=a$, via solutions of (5), into terminal values, (B', n'), at $x=b$. This is a homeomorphism which maps the line, $A'=A$, in the (A', m')-plane into some curve, Γ, bounded by the two lines, $B'=\underline{B}$ and $B'=\bar{B}$, in the (B', n')-plane.

Now it will be shown that Γ cannot be bounded above or below in the n' direction. For if Γ were bounded above by the line $n'=M$, for example, then since the region outside the semi-infinite rectangle containing Γ formed by the lines, $B'=\underline{B}$, $n'=M$, and $B'=\bar{B}$, is arcwise connected the same must be true of its image under the inverse mapping, T^{-1}. Hence this image must lie completely to one side of $A'=A$, and the previous assumption on A for B assures that the image lies to the left of $A'=A$. Consequently the images, Γ_j, for $j=1, 2, 3, \cdots$, of the lines $A'=A+1, A+2, A+3, \cdots$, all lie inside the semi-infinite rectangle containing Γ.

Let γ_j be the supremum in the n' direction of Γ_j and $\gamma_J^*=\min_{j\leqq J}\gamma_j$. Since the norm of a solution at a is unbounded for the family corresponding to the line, $A'=A+j$, the norm of such solutions at b must be similarly unbounded. Thus, for $j\leqq J$, Γ_j must intersect the line, $n'=\gamma_J^*-1$, and at least two of these Γ_j's must cut this line at points, (B_{J1}, γ_J^*-1) and (B_{J2}, γ_J^*-1), with $|B_{J2}-B_{J1}|\leqq(\bar{B}-\underline{B})/(J-1)$, which goes to 0 as $J\to\infty$. Thus the norm at $x=b$ of the difference of the solutions of (5) corresponding to (B_{J1}, γ_J^*-1) and (B_{J2}, γ_J^*-1) goes to 0 as $J\to\infty$, and by Lemma 1 (with $\varphi=\psi=0$) the norm at $x=a$ of this difference must go to 0 as $J\to\infty$. This last is impossible since the norm at $x=a$ of any such difference is at least 1 for two solutions corresponding to points on $A'=A+j_1$, and $A'=A+j_2$, respectively, with $j_1\neq j_2$. This contradiction shows that Γ is not bounded above nor below in the n' direction.

Since the image under T of the line, $A'=A$, is a curve, Γ, lying between the lines, $B'=\underline{B}$ and $B'=\bar{B}$, and unbounded above and below, T^{-1} must map points on opposite sides of the strip containing Γ into points on opposite sides of $A'=A$. In view of our original assumption on A for B, it follows that all solutions through (b, B^*) with $B^*<\underline{B}$ must be bounded below by A at $x=a$. By Theorem 1, for any fixed B^{**} all solutions through (b, B^{**}) must be bounded above and below at $x=a$.

Finally for y a solution through (a, A), let z be the solution with $z(b)=\underline{B}\leqq y(b)$ and $z'(b)=y'(b)$. Were $z(c)>y(c)$ for some c with $|c-b|$ sufficiently small to guarantee uniqueness of solutions to second boundary value problems, it would be a contradiction. Hence $z\leqq y$ on an interval, I, about b. Since Γ is not bounded above or below, the set of $y'(b)$'s, hence of $z'(b)$'s, must be unbounded above and below. By the remark above (9), at any point, c, with $|c-b|$ sufficiently small the set of $z(c)$'s must be unbounded above, and for c also in I the set of $y(c)$'s must be unbounded above. Thus upper existence holds in some deleted neighborhood of b, and similarly lower existence holds in some deleted neighborhood of b. Then b is an isolated point of failure of existence.

Part (iii) follows immediately from (i) and (ii) when we reverse roles of a and b. □

Theorem 3(iii) actually need not hold when

$$y'(a)=m \tag{14'}$$

is used in place of (14), even though Theorem 3 (i), (ii) does. Example 1 provides an instance of failure of Theorem 3 (iii) when the zeros of C on (c,∞) are considered initial points and a is considered as terminal point.

Theorem 2 shows that the three sets of points at which upper existence, lower existence, or existence fail, respectively, consist of isolated points of the set and intervals. Indeed the details of the proof of Theorem 2 show that two components of one of the first two sets are separated by a positive distance depending on the Lipschitz constants valid over a suitably large interval. Theorem 3 shows that the intervals of the first set are disjoint from those of the second. The next theorem shows that the isolated points of the first two sets are also isolated in the third set.

THEOREM 4. *An isolated point of the set at which upper (lower) existence fails is an isolated point of the set at which existence fails.*

Proof. Let b be an isolated point of the set at which upper existence fails. If there exists M such that $y'(b)\leqq M$ for every solution, y, of (5) through (a, A), let z be the solution of (5) such that $z(b)=\bar{B}(b)\geqq y(b)$ and $z'(b)=M\geqq y'(b)$. Suppose $z(c)<y(c)$ for some c in $(b, c_0]$ with $c_0>b$ and c_0 sufficiently close to b to guarantee uniqueness of solution to second boundary value problems on subintervals of $[b, c_0]$. Then there exists c' in $[b, c]$ with $y(c')=z(c')$, $y>z$ on (c', c), $y'(c')\geqq z'(c')$. Thus there exists c'' in $[b, c']$ with $y'(c'')=z'(c'')$, a contradiction of uniqueness of solution of second boundary value problems. Thus $y\leqq z$ on $[b, c_0]$, and upper existence fails on $[b, c_0]$. Since the latter contradicts the isolated character of b, it follows that no such M exists and the $y'(b)$'s are not bounded above. A similar argument assures that they are not bounded below. The argument used in the proof of Theorem 3 (ii) now yields that lower existence also holds in some deleted neighborhood of b. Thus b is an isolated point of the set where existence fails. □

The results on sets of points of existence failure are readily translated into terms of existence holding. For any point b, Theorem 2 yields an open nonvoid interval adjacent to b on right (and similarly on left) in which existence holds everywhere or in which upper existence fails everywhere or in which lower existence fails everywhere, and Theorem 3 assures that at most one of the failure types occurs at any point of the interval. If, in addition, upper (lower) existence holds at b, then the procedure of Theorem 2 assures that upper (lower) existence cannot fail on both the left and right intervals; so b is in a nondegenerate interval for which upper (lower) existence holds. (Actually in this case a minimum length for a containing interval can be determined from the Lipschitz constants.) Thus the set of points where upper (lower) existence holds consists of nondegenerate intervals. Then the set of points where existence holds consists of intervals, possibly degenerating to points. The component intervals of the set for which some specific existence condition holds are isolated in the sense that Theorem 2 shows no finite interval contains infinitely many of them.

Example 2. In the case $f(x, y, y') \equiv |y|$, $\varphi(x) \equiv 0$, $a = 0$, $A = -1$, one can easily compute the family of solutions and find that upper existence fails on the interval $[\pi, +\infty)$ but the upper bounding curve now is the *envelope* of the family of solutions which have the following form where they are nonnegative, with the relevant values of α for each x indicated:

$$y = \sin(x-\alpha)\operatorname{csch}\alpha, \qquad x-\pi \leqq \alpha \leqq x, \quad \alpha > 0.$$

This envelope is easily found to be (in parametric form)

$$x = \pi + \operatorname{arctanh}(1/m) - \arctan(1/m)$$
$$y = \sqrt{\frac{m^2-1}{m^2+1}}, \qquad m > 1,$$

which is a curve passing through the point $(\pi, 1)$ and decreasing monotonely to zero at $+\infty$. The slope of this upper bounding curve $\bar{B}(x)$ tends to $-\infty$ as $x \to \pi$ from the right. (See Fig. 1.)

Example 3.

$$f(x, y, y') \equiv \begin{cases} y & \text{if } y \geqq 0 \\ 4y & \text{if } y \leqq 0 \end{cases} \qquad a = 0, \quad A = 1.$$

Obviously the family of solutions can again be found easily, and we obtain the result that lower existence fails on the interval $[\pi/2, \pi]$ (as well as on others) and that the lower bounding curve there is again found to be simply the envelope of the family of solutions which have the following form where they are nonpositive on $[\pi/2, \pi]$, with the relevant values of α for each x indicated:

$$y(x) = \tfrac{1}{2}\sin(2x-2\alpha)\sec\alpha, \qquad x-\pi \leqq \alpha \leqq x - \frac{\pi}{2}.$$

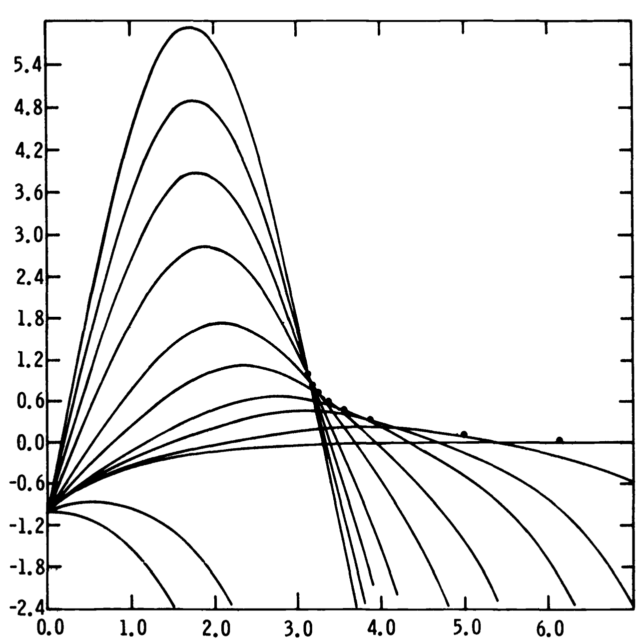

FIG. 1. *The solid lines indicate some of the solutions of the equation of Example 2, and the small circles lie along the envelope.*

This envelope is easily found to be (in parametric form)

$$x = 3\pi/4 - \arctan m + \tfrac{1}{2}\arctan(\tfrac{1}{2}m)$$

$$y = -\tfrac{1}{2}\cos(\arctan(\tfrac{1}{2}m))\sec(\arctan m).$$

In this case the lower bounding curve $y = \underset{=}{B}(x)$ is a symmetrical curve about the vertical line $x = 3\pi/4$, passing through $(\pi/2, -1)$, $(3\pi/4, -\tfrac{1}{2})$, $(\pi, -1)$, with infinite slope at each end of the interval. (See Fig. 2.)

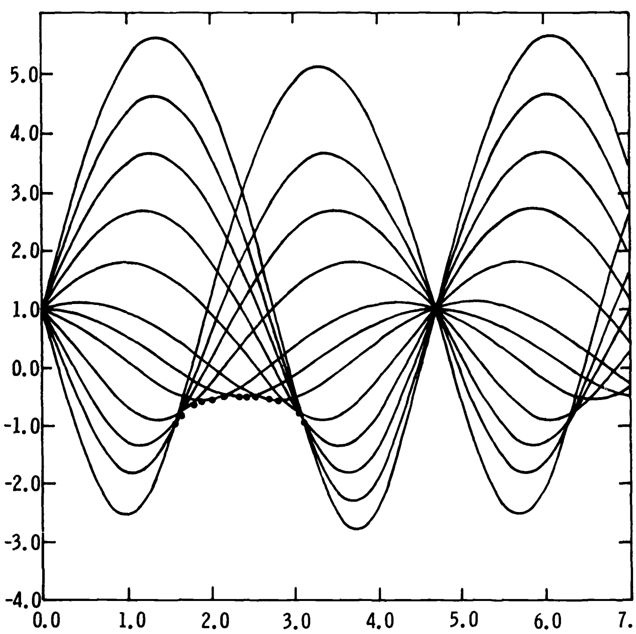

FIG. 2. *The solid lines indicate some of the solutions of the equation of Example* 3, *and the small circles lie along the envelope.*

Example 4. This example was constructed to show that an upper or lower bounding curve could fail to have a first derivative at an interior point. Let

$$f(x, y, y') \equiv \begin{cases} y+1 & \text{if } y \geqq 3, \\ 2y-2 & \text{if } 2 \leqq y \leqq 3, \\ y & \text{if } 0 \leqq y \leqq 2, \\ 4y & \text{if } y \leqq 0, \end{cases}$$

and take $a = 0$, $A = -1$. Some of the solutions, $y(x; a, A, m)$, of (5) through (a, A) are shown in Fig. 3. The small circles in the middle of the figure lie on the envelope of the family of solutions. In the upper right-hand corner of the figure a duplicate copy of the envelope has been inserted so that its form may be seen more clearly. In this example one can see that the envelope has a loop, which is necessarily missing from the lower bounding curve. This is an obvious consequence of the definition of lower bounding curve and in this example at least is clearly the cause of the lack of a first derivative of the lower bounding curve.

The complete analysis for the above example and some slightly simpler ones has been carried out, but the details are rather tedious and are omitted.

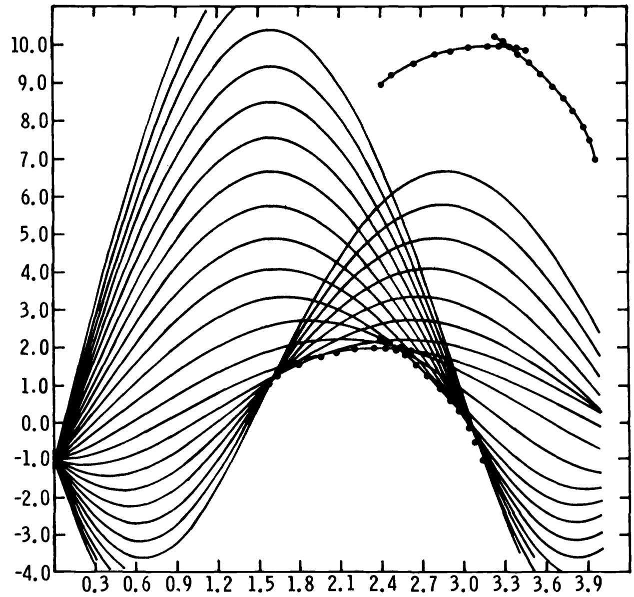

FIG. 3. *The solid lines indicate some of the solutions of the equation of Example 4, and the small circles lie along the envelope. For greater clarity the envelope has been duplicated in the upper righthand corner. Evidently the lower bounding curve in this case is not differentiable at the point where the envelope crosses itself.*

As one can see by looking at Example 4, the upper or lower bounding curves need not be differentiable; but where they are, it turns out that they are envelopes of the family of solutions in the sense of Theorem 5 below. (They may be actual solutions of the differential equation in some cases. See example in (15).)

THEOREM 5. *Let upper existence fail for b on the interval* $[b_1, b_2]$, *with* $b_1 < b_2$, *for the family, S, of solutions of* (5) *and* (14). *If* $\bar{B}$ *is the bounding function on* $[b_1, b_2]$, *and, for b in* $[b_1, b_2]$, S_b *is the subfamily of S consisting of solutions which take value,* $\bar{B}(b)$, *at b; then*

(i) $\bar{B}$ *is continuous,*

(ii) *for b in* (b_1, b_2), S_b *is not empty,*

(iii) $\bar{B}$ *has right and left derivatives,* $\bar{B}_R$ *and* $\bar{B}_L$, *on* (b_1, b_2).

For b in (b_1, b_2), $\bar{B}_L(b) \leqq y'(b) \leqq \bar{B}_R(b)$ *for every y in* S_b, *and there exist elements of* S_b *for which the equalities are attained.* $\bar{B}$ *has a derivative at b in* (b_1, b_2) *if and only if* S_b *consists of a single function. If* S_{b_1} *is empty,* $\bar{B}$ *has right differential coefficient of* $-\infty$ *at* b_1; *and if* S_{b_2} *is empty,* $\bar{B}$ *has left differential coefficient of* ∞ *at* b_2. *If* S_{b_1} *or* S_{b_2} *is not empty, the corresponding one-sided differentiation is as described for interior points.*

(iv) *For b in* (b_1, b_2)

$$\lim_{c\to b-} \bar{B}_L(c) = \lim_{c\to b-} \bar{B}_R(c) = \bar{B}_L(b) \leqq \bar{B}_R(b) = \lim_{c\to b+} \bar{B}_L(c) = \lim_{c\to b+} \bar{B}_R(c).$$

For $b = b_1$ *or* $b = b_2$ *and* S_{b_i} *not empty, the appropriate one-sided results hold. For* S_{b_i} *empty, the limits, in the sense of proper divergence, are* $-\infty$ *or* ∞ *as appropriate.*

$\bar{B}_L$ *and* $\bar{B}_R$ *are bounded on any closed subinterval of* (b_1, b_2). $\bar{B}_L$ *and* $\bar{B}_R$ *have the same points of discontinuity in* (b_1, b_2), *specifically the points at which they are unequal (i.e.,* $\{b \mid (\text{cardinal of } S_b) > 1\}$). *This set of discontinuities is countable.*

(v) *The lower derivatives,* $\bar{B}_{R,-}$ *and* $\bar{B}_{L,-}$, *of* $\bar{B}_R$ *and* $\bar{B}_L$, *respectively, satisfy*

$$\bar{B}_{R,-}(x) + f(x, \bar{B}(x), \bar{B}_R(x)) \geqq 0$$

$$\bar{B}_{L,-}(x) + f(x, \bar{B}(x), \bar{B}_L(x)) \geqq 0$$

on (b_1, b_2). (*The inequalities also hold at* b_1 *and* b_2 *if meaningful when the second derivative is taken as one-sided.*) *For any closed subinterval, I, of* (b_1, b_2) *there exists a constant,* M_I, *such that* $\bar{B}_R(x)+M_Ix$ *and* $\bar{B}_L(x)+M_Ix$ *are increasing on I.* $\bar{B}_R$ *and* $\bar{B}_L$ *have equal derivatives almost everywhere on* (b_1, b_2).

In the case of lower existence failure on $[b_1, b_2]$ *an exactly corresponding set of results holds.*

Proof. There is a positive number such that first and second boundary value problems have unique solutions on any subinterval of $[b_1, b_2]$ with length at most this number. To simplify statements of arguments let $K(x)$ be the intersection of $[b_1, b_2]$ and the closed interval with this length centered at x. Note that, for x_1, x_2 in $[b_1, b_2]$, x_1 is in $K(x_2)$ if and only if x_2 is in $K(x_1)$.

(i) Given b in $[b_1, b_2)$ and $\varepsilon > 0$, take y_1 in S so that $y_1(b) \geqq \bar{B}(b)-\varepsilon/2$. Then for h positive and sufficiently small $\bar{B}(b+h) \geqq y_1(b+h) \geqq y_1(b)-\varepsilon/2 \geqq \bar{B}(b)-\varepsilon$.

For some positive h' with $b+h'$ in $K(b)$ take y_2 as a solution of (5) with $y_2(b)=\bar{B}(b)$ and $y_2(b+h')=\bar{B}(b+h')$. As was seen in the proof of Theorem 2, uniqueness of solution to FBVP assures $\bar{B} \leqq y_2$ on $[b, b+h']$. For h positive and sufficiently small $\bar{B}(b+h) \leqq y_2(b+h) \leqq y_2(b)+\varepsilon = \bar{B}(b)+\varepsilon$. Thus $\bar{B}$ is continuous from the right at b.

Similarly for b in $(b_1, b_2]$, $\bar{B}$ is continuous from the left at b. Thus $\bar{B}$ is continuous on $[b_1, b_2]$.

(ii) Let $b \in (b_1, b_2)$. For each positive integer, n, choose y_n in S such that $y_n(b) \geqq \bar{B}(b)-1/n$. Let $\bar{y}_n$ be the solution of (5) such that $\bar{y}_n(b)=\bar{B}(b)-1 \leqq y_n(b)$, $\bar{y}_n'(b)=y_n'(b)$. Uniqueness of solution for SBVP assures $\bar{y}_n \leqq y_n \leqq \bar{B}$ on $K(b)$. By the remark above (9), in order to have $\bar{y}_n$'s bounded above by $\bar{B}$ on $K(b)$ it is necessary that the $\bar{y}_n'(b)$'s, that is, the $y_n'(b)$'s, be bounded. By choosing an appropriate subsequence, we can assume without loss of generality that $\{y_n'(b)\}$ converges. Thus $\{y_n\}$ and $\{y_n'\}$ converge to y and y' where y is a solution of (5) satisfying the boundary condition at a. Since $y(b)=\bar{B}$, $y \in S_b$. Thus S_b is not empty.

(iii) For c_1, c_2 in (b_1, b_2) with $c_1 < c_2$, let $y_i \in S_{c_i}$, $i=1, 2$. Since $y_2(c_1) \leqq y_1(c_1)$ and $y_1(c_2) \leqq y_2(c_2)$, there is a c_3 in $[c_1, c_2]$ with $y_1(c_3)=y_2(c_3)$. Uniqueness of solution to FBVP assures $y_2 \leqq y_1$ to left of c_3 on $K(c_3)$ and $y_1 \leqq y_2$ to right of c_3 on $K(c_3)$. Thus $y_1'(c_3) \leqq y_2'(c_3)$; and uniqueness of solution to SBVP assures that $y_1' \leqq y_2'$ on $K(c_3)$.

Note that the above argument would still hold even though c_i were permitted to be b_i provided S_{b_i} were not empty.

For c in (b_1, b_2) let y_c be an arbitrary element of S_c. (There might be many elements in some of the S_c's; so there might be many choices for this function mapping (b_1, b_2) into S, and any specific function is taken.)

In the following discussion in (iii) of right differentiation, $b \in [b_1, b_2)$ and abscissas are restricted to be in $K(b)$ and not less than b unless otherwise noted. Abscissas specifically introduced by name, other than b itself, are to be greater than b.

If $c_1 \leqq c_2$, then $b \in K(c_3)$ and is to left of c_3 for any c_3 in $[c_1, c_2]$ and, by result above, one has $y_{c_2}(b) \leq y_{c_1}(b)$ and $y_{c_1}'(b) \leq y_{c_2}'(b)$. Thus $y_c(b)$ is decreasing in c and bounded above by $\bar{B}(b)$ and $y_c'(b)$ is increasing in c. In addition, unless b is b_1, the $y_c'(b)$'s are bounded below by $y_b'(b)$.

Case 1. *b is* b_1 *and* $y_c'(b_1)$*'s are not bounded below.* Let z_c be the solution of (5) with $z_c(b_1)=\bar{B}(b_1) \geqq y_c(b_1)$ and $z_c'(b_1)=y_c'(b_1)$. Then for $c_1 \leqq c_2$, the standard arguments shows $z_{c_2} \geqq z_{c_1} \geqq y_{c_1}$. Thus

$$\bar{B}(c_1)-\bar{B}(b_1)=y_{c_1}(c_1)-z_{c_2}(b_1) \leqq z_{c_2}(c_1)-z_{c_2}(b_1),$$

and the upper right differential coefficient of $\bar{B}$ at b_1 is at most $z_{c_2}'(b_1)$, or $y_{c_2}'(b_1)$. Thus the upper right differential coefficient of $\bar{B}$ at b_1 is $-\infty$.

Case 2. *$y_c'(b)$'s are bounded below, thus covering the case that b is not b_1.* Now $\{y_c(b)\}$ and $\{y_c'(b)\}$ converge as c approaches b. Thus on any finite interval (not just $K(b)$), $\{y_c\}$ and $\{y_c'\}$ converge uniformly to y and y' as c approaches b, where y is a solution of (5) satisfying the boundary condition at a.

If $y(b)<\bar{B}(b)$, then there exist ε and δ, both positive, such that $y \leqq \bar{B}-2\varepsilon$ on $[b, b+\delta]$. Thus, for c sufficiently close to b, $y_c \leqq \bar{B}-\varepsilon$ on $[b, b+\delta]$, a contradiction since $y_c(c)=\bar{B}(c)$. Thus $y(b)=\bar{B}(b)$ and $y \in S_b$.

Now

$$\frac{\bar{B}(c)-\bar{B}(b)}{c-b} \leqq \frac{y_c(c)-y_c(b)}{c-b} = y_c'(b+\theta_{c1}(c-b)), \qquad 0<\theta_{c1}<1,$$

$$\frac{\bar{B}(c)-\bar{B}(b)}{c-b} \geqq \frac{y(c)-y(b)}{c-b} = y'(b+\theta_{c2}(c-b)), \qquad 0<\theta_{c2}<1.$$

Given $\varepsilon>0$ there exists a positive δ such that $c \leqq b+\delta$ yields

$$|y_c'-y'| \leqq \varepsilon/2$$

$$|y'(c)-y'(b)| \leqq \varepsilon/2.$$

Thus for $c \leqq b+\delta$,

$$\begin{aligned}|y_c'(b+\theta_{c1}(c-b))-y'(b)| &\leqq |y_c'(b+\theta_{c1}(c-b))-y'(b+\theta_{c1}(c-b))| \\ &\quad +|y'(b+\theta_{c1}(c-b))-y'(b)| \\ &\leqq \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,\end{aligned}$$

$$|y'(b+\theta_{c2}(c-b))-y'(b)| \leqq \frac{\varepsilon}{2}.$$

Hence $\bar{B}$ has right derivative, $\bar{B}_R(b)$, at b equal to $y'(b)$. For z in S_b, $\bar{B} \geqq z$ and $\bar{B}(b)=z(b)$. Thus $\bar{B}_R(b) \geqq z'(b)$, and we have already shown that equality holds for $z=y$.

Finally note that if b is b_1, existence of a lower bound for the $y_c'(b_1)$'s guarantees that S_{b_1} is not empty.

The corresponding results for left differentiation are obtained analogously. On (b_1, b_2), $\bar{B}$ has a derivative if and only if $\bar{B}_L=\bar{B}_R$, thus if and only if S_b consists of a single function.

(iv) By (iii), at any c in (b_1, b_2), either right derivative, $\bar{B}_R(c)$, or left derivative, $\bar{B}_L(c)$, can be replaced by $y_c'(c)$ where y_c is chosen suitably from S_c. If $b \in (b_1, b_2)$ and $c \in (b, b_2)$, it has been seen that $\{y_c'\}$ converges uniformly on any finite interval to y' as $c \to b+$, *where* $y'(b)=\bar{B}_R(b)$. Thus

$$\bar{B}_{\binom{R}{L}}(c)-\bar{B}_R(b)=y_c'(c)-y'(b)=(y_c'(c)-y'(c))+(y'(c)-y'(b)),$$

and since both of the terms on the right may be made as small in magnitude as desired by restricting c to be sufficiently close to b,

$$\lim \bar{B}_{\binom{R}{L}}(c)=\bar{B}_R(b).$$

If there existed a sequence, $\{c_n\}$, converging monotonely to b_1 from the right for which $\{y'_{c_n}(c_n)\}$ was bounded below, then since $y'_{c_n}(c_n) \leqq y'_{c_k}(c_n)$ for c_k in $K(b_1)$ and $n \geqq k$, $\{y'_{c_n}(c_n)\}$ would be bounded above. By extracting a convergent subsequence and re-indexing one may assume that $\{y'_{c_n}(c_n)\}$ converges to m, say. Then $\{(c_n, y_{c_n}(c_n), y'_{c_n}(c_n))\}$ is equal to $\{(c_n, \bar{B}(c_n), y'_{c_n}(c_n))\}$, which converges to $(b_1, \bar{B}(b_1), m)$. Thus $\{y_{c_n}\}$ would converge as usual to a solution, y, in S_{b_1}, and S_{b_1} would not be empty.

If S_{b_1} is empty, it follows from the above that

$$\lim_{c \to b_1+} y'_c(c) = -\infty$$

that is,

$$\lim_{c \to b_1+} \bar{B}_{\binom{R}{L}}(c) = -\infty.$$

If S_{b_1} is not empty, the argument used for interior points holds.

It is easy to see that, on any closed subinterval of (b_1, b_2), $\bar{B}_R$ attains its maximum and $\bar{B}_L$ its minimum. Since $\bar{B}_L \leqq \bar{B}_R$, both are bounded on the interval.

Plainly from the established results $\bar{B}_L$ and $\bar{B}_R$ are continuous at b if $\bar{B}_L(b) = \bar{B}_R(b)$, and each has a positive jump discontinuity from the appropriate side if $\bar{B}_L(b) \neq \bar{B}_R(b)$. It is known [3, p. 392] that the set of points at which the right derivative exceeds the left derivative is countable. (That the set of discontinuities is countable in the present situation may be verified in other ways, a slightly stronger result being a consequence of (v).)

(v) For c in (b_1, b_2), let $y_c \in S_c$ with $y'_c(c) = \bar{B}_R(c)$. Then if b, $c \in (b_1, b_2)$ and $c \in K(b)$,

$$\frac{\bar{B}_R(c) - \bar{B}_R(b)}{c-b} = \frac{y'_c(c) - y'_b(b)}{c-b} = \frac{y'_c(c) - y'_b(c)}{c-b} + \frac{y'_b(c) - y'_b(b)}{c-b} \geqq \frac{y'_b(c) - y'_b(b)}{c-b}$$

since $y'_c - y'_b$ has the same sign as $c-b$ on $K(c_3)$ for some c_3 between b and c. Thus the lower derivative, $\bar{B}_{R,-}(b)$, of $\bar{B}_R$ at b satisfies

$$\bar{B}_{R,-}(b) \geqq y''_b(b),$$

$$\begin{aligned}\bar{B}_{R,-}(b) + f(b, \bar{B}(b), \bar{B}_R(b)) &\geqq y''_b(b) + f(b, \bar{B}(b), \bar{B}_R(b)) \\ &= y''_b(b) + f(b, y_b(b), y'_b(b)) = 0.\end{aligned}$$

If S_{b_1} is not empty, the argument holds at b_1 provided the lower derivative of $\bar{B}_R$ at b_1 is considered as taken from the right.

If I is a closed subinterval of (b_1, b_2), $\bar{B}_R(x)$ is bounded on I and so is $|f(x, \bar{B}(x), \bar{B}_R(x))|$, say by $(M_I - 1)$ for the latter. Then $\bar{B}_R(x) + M_I x$ has lower derivative on I which is positive, so $\bar{B}_R(x) + M_I x$ is increasing there.

The analogous results hold for $\bar{B}_L$, with a common M_I being chosen for both.

As increasing functions, $\bar{B}_R(x) + M_I x$ and $\bar{B}_L(x) + M_I x$ have derivatives almost everywhere, and thus so do $\bar{B}_R$ and $\bar{B}_L$. The established relations between $\bar{B}_R$ and $\bar{B}_L$ show that both can have derivatives only at points where $\bar{B}_R = \bar{B}_L$ and that the derivatives are equal when both exist.

In the case of lower existence failure on $[b_1, b_2]$ the corresponding results can obviously be proved in exactly the same way. □

While it is of course true that each solution of (5) is bounded on every finite interval, an upper or lower bounding curve of the family of solutions passing through a given point (a, A) may very well be unbounded, as the following example shows.

Example 5.

$$y''+f(y)=0, \qquad y(0)=1, \tag{16}$$

$$y(b)=B, \tag{17}$$

where

$$f(y)=\begin{cases} y-2\sqrt{y} & \text{for } y\geqq 1 \\ -y & \text{for } y\leqq 1. \end{cases} \tag{18}$$

Clearly f is continuous and Lipschitzian, since $-1\leqq f'(y)\leqq 1$ on each half-line. We shall show that upper existence fails for $b>\pi$, and that $\bar{B}(x)$ tends to $+\infty$ as $x\to\pi+$.

Notice first that if $v(x)$ is defined by

$$v''=-v+2, \quad v(0)=1, \quad v'(0)=m>0,$$

then

$$v(x)=2-\cos x+m\sin x,$$

so $v(\pi)=3$. Let $y(x; m)$ denote the solution of (16) which agrees with v in value and slope at $x=0$. Then $y(x; m)$ satisfies

$$y''\geqq -y+2$$

(at least as long as $y\geqq 1$) so by well known comparison theorems (for example [2, p. 80])

$$y(x; m)\geqq v(x) \quad \text{on } [0, \pi).$$

In particular, $y(\pi; m)\geqq 3$ for $m>0$.

Now since

$$\frac{1}{2}\frac{d}{dy}(y')^2+f(y)=0,$$

then

$$\frac{dy}{dx}=\pm\left[m^2-y^2+\frac{8}{3}y^{3/2}-\frac{5}{3}\right]^{1/2}.$$

Denote by X the first value of $x>0$ for which $dy/dx=0$, and let the corresponding maximum value of y be Y. (Both are functions of m.) Then

$$X=\int_1^Y\left[m^2-\frac{5}{3}-z^2+\frac{8}{3}z^{3/2}\right]^{-1/2}dz,$$

and for $X\leqq x\leqq$ (first zero of $(y-1)$ after X)

$$x=X+\int_y^Y\left[m^2-\frac{5}{3}-z^2+\frac{8}{3}z^{3/2}\right]^{-1/2}dz. \tag{19}$$

By definition of Y,

$$m^2-\frac{5}{3}-Y^2+\frac{8}{3}Y^{3/2}=0,$$

so

$$Y \to \infty \qquad \text{as } m \to \infty.$$

Also, for $1 \leqq y \leqq Y$,

$$\int_y^Y \left[m^2 - \frac{5}{3} - z^2 + \frac{8}{3} z^{3/2}\right]^{-1/2} dz = \int_y^Y \left[Y^2 - z^2 - \frac{8}{3}(Y^{3/2} - z^{3/2})\right]^{-1/2} dz$$
$$= \int_{y/Y}^1 \left[1 - t^2 - \frac{8}{3Y^{1/2}}(1 - t^{3/2})\right]^{-1/2} dt.$$

In particular,

$$X = \int_{1/Y}^1 \left[1 - t^2 - \frac{8}{3Y^{1/2}}(1 - t^{3/2})\right]^{-1/2} dt \to \int_0^1 (1 - t^2)^{-1/2} dt = \frac{\pi}{2} \quad \text{as } m \to \infty.$$

By symmetry from (19), when $y(x; m)$ is next equal to 1, $x = 2X$, which tends to π as $m \to +\infty$. Thus for large positive m, $y(x; m)$ rises to a maximum of Y at X, then drops to the value of 1 again at $x = 2X$, which is nearly π, where it has slope $-m$. Such a solution necessarily has a zero quite close to $x = \pi$, and must remain negative thereafter in view of the definition of f. From this it follows that upper existence fails for every $b > \pi$.

Lastly, let us show that $y(\pi; m) \to +\infty$ as $m \to +\infty$, since this will imply that $\bar{B}(x) \to +\infty$ as $x \to \pi+$. Since (19) can be written in the form

$$x = \int_u^1 2t|1 - v - t^4 + vt^3|^{-1/2} dt + \int_{uy^{1/2}}^1 2t|1 - v - t^4 + vt^3|^{-1/2} dt \tag{20}$$

with $u = Y^{-1/2}$ and $v = 8u/3$, it will be sufficient to show that the solution $y(u)$ of the equation

$$\pi = \int_u^1 2t|1 - v - t^4 + vt^3|^{-1/2} dt + \int_{uy^{1/2}}^1 2t|1 - v - t^4 + vt^3|^{-1/2} dt \tag{21}$$

is not bounded as $u \to 0$.

Using the relation

$$\int_u^1 2t(1 - t^4)^{-1/2} dt = \pi/2 - \arcsin u^2,$$

we can write (21) as

$$\arcsin u^2 + \int_u^{uy^{1/2}} t[1 - v - t^4 + vt^3]^{-1/2} dt$$
$$= 2 \int_u^1 t\{[1 - v - t^4 + vt^3]^{-1/2} - (1 - t^4)^{-1/2}\} dt,$$

or

$$u^{-1} \arcsin u^2 + u \int_1^{y^{1/2}} s[1 - v - u^4 s^4 + vu^3 s^3]^{-1/2} ds$$
$$= \frac{16}{3} \int_u^1 t(1 - t^3)(1 - t^4)^{-1/2}[1 - v - t^4 + vt^3]^{-1/2} \cdot \{(1 - t^4)^{1/2} + [1 - v - t^4 + vt^3]^{1/2}\}^{-1} dt. \tag{22}$$

Obviously the right-hand side of this equation has a positive limit as $u \to 0$. But if $y(u)$ is bounded as $u \to 0$, the limit of the left hand side is zero, which is a contradiction. Thus $y(u)$ must be unbounded as claimed. □

5. Dependence on initial conditions and reciprocity. Plainly, if b is any point such that for solutions from $x = a$ existence fails at $x = b$, then, vice versa, for solutions from $x = b$ existence fails at $x = a$. Theorem 3 shows that in this situation if both upper and lower existence fail from a to b, then both upper and lower existence fail from b to a. Moreover, when the boundary condition at a is (14), both cases are of existence failure at an isolated point.

While such reciprocity, when the roles of initial point and terminal point are reversed, has not been established for all possible types of existence failure, it does hold also in the case that b is in the interior of the set where upper (lower) existence fails; that is, when b is considered as initial point then a is in the interior of one of the sets of one-sided existence failure.

The result is a consequence of the sequence of theorems which follow providing information on the behavior of the bounding curves with respect to initial conditions.

THEOREM 6. *If there exists an open interval, I_0, about b such that upper (lower) existence fails from a to b' whenever $b' \in I_0$, then there exists a subinterval, I, about b and an interval, J, about a such that upper (lower) existence fails from a' to b' whenever $a' \in J$ and $b' \in I$.*

The proof of this theorem consists essentially of showing that although the difference at $x = b$ between the solutions through (a, A) and (a', A), respectively, both with derivative, m, at the initial point may tend to $\pm\infty$ as $|m| \to \infty$, the hypotheses assure that the solutions through (a, A) have values at $x = b$ which tend to $-\infty$ even more rapidly as $|m| \to \infty$, provided $|a' - a|$ is sufficiently small.

The next two lemmas provide suitable estimates of the rates of growth of solutions to both initial value problems and boundary value problems with respect to a given slope.

LEMMA 2. *If y is a solution of* (5), *then*

$$|y(x) - y(a) - y'(a)(x-a)| \leqq [D_x + K|y(a)| + (K|x-a| + L)|y'(a)|]k^{-1} \exp k|x-a| - 1)|x-a| \tag{23}$$

where $D_x = \max_{x' \in [a,x]} |f(x', 0, 0)|$, K and L are the Lipschitz constants in (6), *and $k = \max\{K, L+1\}$.*

Proof. For use in Lemma 1 take $u(x) = y(x)$ and $v(x) = y(a) + y'(a)(x-a)$. Then $v''(x) + f(x, v(x), v'(x)) = f(x, y(a) + y'(a)(x-a), y'(a))$, $v(a) = u(a)$, and $v'(a) = u'(a)$. Thus $\|u - v\|(a) = 0$, $\varphi(x) = 0$, and $\psi(x) = f(x, y(a) + y'(a)(x-a), y'(a))$. Thus, with $M = 0$, (12) gives

$$\|y - v\|(x) \leqq Nk^{-1}(\exp k|x-a| - 1). \tag{24}$$

Now

$$\begin{aligned} N &= \max_{x' \in [a,x]} |f(x', y(a) + y'(a)(x'-a), y'(a))| \\ &\leqq \max_{x' \in [a,x]} [|f(x', 0, 0)| + |f(x', y(a) + y'(a)(x'-a), y'(a)) - f(x', 0, 0)|] \\ &\leqq \max_{x' \in [a,x]} [|f(x', 0, 0)| + K|y(a) + y'(a)(x'-a)| + L|y'(a)|] \\ &\leqq D_x + K|y(a)| + (K|x-a| + L)|y'(a)| \end{aligned} \tag{25}$$

where $D_x = \max_{x' \in [a,x]} |f(x', 0, 0)|$.

Equations (24) and (25) give

$$\|y-v\|(x) \leqq [D_x + K|y(a)| + (K|x-a|+L)|y'(a)|]k^{-1}(\exp k|x-a|-1). \tag{26}$$

In particular, $|y'(x)-y'(a)| = |y'(x)-v'(x)| \leqq \|y-v\|(x)$, and $|y'(x)-y'(a)|$ is bounded by the right side of (26). Integrating from a to b while bounding $|x-a|$ by $|b-a|$ and D_x by D_b yields (23) with $x=b$. □

LEMMA 3. *If y is a solution of* (5), *then*

$$|y(a)-y(b)| \leqq \{(D+K|y(b)|)k^{-1}(E-1) + [(K|b-a|+L)k^{-1}(E-1)+1]|y'(a)|\}|b-a| \div \{1-Kk^{-1}(E-1)|b-a|\} \tag{27}$$

where $D = \max_{x\in[a,b]} |f(x,0,0)|$, $E = \exp k|b-a|$, *and* $|b-a|$ *is sufficiently small that the denominator is positive. Also*

$$|y'(b)-y'(a)| \leqq \frac{[D+K|y(b)|+(K|b-a|+L)|y'(a)|]k^{-1}(E-1)}{[1-(K|b-a|+L)k^{-1}(E-1)]} \tag{28}$$

where $|b-a|$ *is sufficiently small that denominator is positive.*

Proof. Using Lemma 2, with $x=b$, we obtain

$$|y(b)-y(a)-y'(a)(b-a)| \leqq [D+K|y(a)|+(K|b-a|+L)|y'(a)|]k^{-1}(E-1)|b-a|,$$

$$\begin{aligned} |y(a)-y(b)| &\leqq [D+K(|y(a)-y(b)|+|y(b)|)+(K|b-a|+L)|y'(a)|]k^{-1}(E-1)|b-a| \\ &\quad + |y'(a)|\,|b-a| \\ &= \{(D+K|y(b)|)k^{-1}(E-1)+[(K|b-a|+L)k^{-1}(E-1)+1]|y'(a)|\}|b-a| \\ &\quad + Kk^{-1}(E-1)|b-a|\,|y(a)-y(b)|. \end{aligned}$$

Transposing term on right involving $|y(a)-y(b)|$ and dividing yields (27). □

From the remark following (26) in proof of Lemma 2 with a and x replaced by b and a one has

$$\begin{aligned} |y'(a)-y'(b)| &\leqq [D+K|y(b)|+(K|a-b|+L)|y'(b)|]k^{-1}(E-1) \\ &\leqq [D+K|y(b)|+(K|a-b|+L)|y'(a)|]k^{-1}(E-1) \\ &\quad + (K|a-b|+L)k^{-1}(E-1)|y'(b)-y'(a)| \end{aligned}$$

and (28) follows.

Proof of Theorem 6. Let Δ be some finite open interval containing $[a,b]$ (we are assuming $a<b$; a similar proof holds if $b<a$), and let δ_0 be positive and sufficiently small that $[a-\delta_0, b+\delta_0] \subset \Delta$, $[b-\delta_0, b+\delta_0] \subset I_0$, and first and second boundary value problems have unique solutions on all subintervals of $[b-\delta_0, b+\delta_0]$. For simplicity of notation we assume that upper existence fails from a to b' whenever $|b'-b| < \delta_0$.

Let $E = \exp k\delta_0$, $F = \exp k \cdot$ (length of Δ), and $D = \max_{x\in\Delta} |f(x,0,0)|$.

For usage in later proofs the development is carried out for an arbitrary number, ε, in $(0,1)$. The fact that $\varepsilon<1$ is used in verifying some of the inequalities. Also the reader should note that when a compact set rather than a single point, b, is being considered if Δ and δ_0 have been chosen so that the conditions of the first paragraph are met for all points of the set, then the choice of the quantities, E, F, D, $\delta(\varepsilon)$, C_1, $G(\varepsilon)$, C_2, and $H(\varepsilon)$, can be made uniformly for all points of the set. Also neither K nor L has been taken as 0.

Define

$$\delta(\varepsilon)=\min\left\{\frac{1}{2}, \delta_0, \frac{\varepsilon}{1+\varepsilon}[K\delta_0+L+k]^{-1}\right\}. \tag{29}$$

It is now easy to verify that $0<d\leqq\delta(\varepsilon)$ assures $kd<1$ and

$$\begin{matrix}(1+\varepsilon)^{-1}<1-(Kd+L)k^{-1}(e^{kd}-1)<1 \\ 1<1+(Kd+L)k^{-1}(e^{kd}-1)<1+\varepsilon\end{matrix} \qquad (0<d\leqq\delta(\varepsilon)). \tag{30}$$

It is convenient to study the behavior of a solution, y, of (5) with $y(a')=A$ by considering the solution of (5) having A as value at a and $y'(a)$ as derivative at a. For comparison purposes the difference between A and $y(a)$ is needed, and Lemma 3 with $b=a'$ and $|a'-a|<\delta(\varepsilon)$ gives

$$|y(a)-A|\leqq\frac{\{(D+K|A|)k^{-1}(\exp k|a'-a|-1)+[(K|a'-a|+L)k^{-1}(\exp k|a'-a|-1)+1]|y'(a)|\}|a'-a|}{\{1-Kk^{-1}(\exp k|a'-a|-1)|a'-a|\}}$$

$$\leqq 2\{(D+K|A|)k^{-1}(E-1)+2|y'(a)|\}|a'-a|.$$

Thus taking

$$C_1=(D+K|A|)k^{-1}(E-1) \tag{31}$$

one has

$$|y(a)-A|\leqq 2(C_1+2|y'(a)|)|a'-a|. \tag{32}$$

(If the boundary condition at a is (14′), then the corresponding condition is $y'(a')=m$. One uses (28) of Lemma 3 with a', a replacing a, b. Then (30) again assures the denominator exceeds $(1+\varepsilon)^{-1}$, $k^{-1}(\exp k|a'-a|-1)$ is a factor, and for fixed m, the remaining factor is bounded by a linear function of $|y(a)|$.)

The next step is to obtain the inequalities involving $u(b)$ and $u'(b)$ for a solution, u, of (5) and (14). Let then $\bar{B}$ be the bounding function on $[b-\delta_0, b+\delta_0]$ and, temporarily, let c be $b+\delta(\varepsilon)$ if $u'(b)\geqq 0$ and be $b-\delta(\varepsilon)$ if $u'(b)<0$. Lemma 2, with b and c in place of a and x and $G(\varepsilon)$ for $(\exp k\delta(\varepsilon)-1)$, yields

$$u(b)+u'(b)(c-b)-\bar{B}(c)\leqq u(b)+u'(b)(c-b)-u(c)\leqq|u(c)-u(b)-u'(b)(c-b)|$$
$$\leqq[D+K|u(b)|+(K\delta(\varepsilon)+L)|u'(b)|]k^{-1}G(\varepsilon)\delta(\varepsilon).$$

Since $u'(b)$ and $c-b$ have same sign,

$$u(b)[1-Kk^{-1}G(\varepsilon)\delta(\varepsilon)\operatorname{sign}u(b)]$$
$$\leqq\bar{B}(c)+Dk^{-1}G(\varepsilon)\delta(\varepsilon)-[1-(K\delta(\varepsilon)+L)k^{-1}G(\varepsilon)]\delta(\varepsilon)|u'(b)|.$$

By (30), both $[1-Kk^{-1}G(\varepsilon)\delta(\varepsilon)\operatorname{sign}u(b)]$ and $[1-(K\delta(\varepsilon)+L)k^{-1}G(\varepsilon)]$ exceed $(1+\varepsilon)^{-1}$, and for $|u'(b)|$ sufficiently large the right side is negative. In such case $u(b)$ is negative, and by (30) one has $[1+Kk^{-1}G(\varepsilon)\delta(\varepsilon)]<1+\varepsilon$. Thus

$$u(b)\leqq\{\bar{B}(c)+Dk^{-1}G(\varepsilon)\delta(\varepsilon)-\tfrac{1}{2}\delta(\varepsilon)|u'(b)|\}/(1+\varepsilon)$$
$$\text{for } |u'(b)|\geqq 2[\bar{B}(c)+Dk^{-1}G(\varepsilon)\delta(\varepsilon)]/\delta(\varepsilon).$$

Letting

$$C_2=Dk^{-1}E \tag{33}$$

one has

$$u(b) \leqq \{\bar{B}(c)+C_2\delta(\varepsilon)-\tfrac{1}{2}|u'(b)|\delta(\varepsilon)\}/(1+\varepsilon) \tag{34}$$
$$\text{for } |u'(b)|>2[\bar{B}(c)+C_2\delta(\varepsilon)]/\delta(\varepsilon).$$

Having established the above result for derivatives at b of large magnitude, one can now use it to show that derivatives at a of large magnitude lead to very negative values at b.

Let u_0 be the solution of (5) and (14) with $u_0'(a)=0$. (If the boundary condition at a is (14′), then take $u_0(a)=0$ and proceed in the same way.) By Lemma 1, with a and b reversed and u a solution of (5) and (14),

$$|u(b)-u_0(b)|+|u'(b)-u_0'(b)|=\|u-u_0\|(b)\geqq\|u-u_0\|(a)e^{-k|b-a|}\geqq F^{-1}|u'(a)|.$$

Since $u(b)-u_0(b)\leqq\bar{B}(b)-u_0(b)$, one could not have $|u'(a)|\geqq 2F(\bar{B}(b)-u_0(b))$ and $u(b)-u_0(b)>\tfrac{1}{2}F^{-1}|u'(a)|$ holding simultaneously. Hence when $|u'(a)|\geqq 2F(\bar{B}(b)-u_0(b))$

$$u(b)\leqq u_0(b)-\tfrac{1}{2}F^{-1}|u'(a)|\leqq\bar{B}(b)-\tfrac{1}{2}F^{-1}|u'(a)| \tag{35i}$$

or

$$|u'(b)-u_0'(b)|\geqq\tfrac{1}{2}F^{-1}|u'(a)|. \tag{35ii}$$

(For either (35ii) holds or else we have both $|u(b)-u_0(b)|>\tfrac{1}{2}F^{-1}|u'(a)|$ and $u(b)-u_0(b)\leqq\tfrac{1}{2}F^{-1}|u'(a)|$, which gives (35i).)

The above pieces are now brought together. Restrict a' by

$$|a'-a|\leqq\delta(\varepsilon)/(32F^2)\leqq\delta(\varepsilon). \tag{36}$$

(For (14′) as boundary condition at a, one must have $k^{-1}(e^{k|a'-a|}-1)\leqq\delta(\varepsilon)/(16F^2K)$.)

Let

$$H(\varepsilon)=2F\max\left(\bar{B}(b)-u_0(b),\,|u_0'(b)|+2\frac{\max(\bar{B}(b-\delta(\varepsilon)),\bar{B}(b+\delta(\varepsilon)))+C_2\delta(\varepsilon)}{\delta(\varepsilon)}\right). \tag{37}$$

For y a solution of (5) with $y(a')=A$, let u be the solution of (5) and (14) with $u'(a)=y'(a)$. Then by Lemma 1 and (32)

$$\begin{aligned} y(b)&\leqq u(b)+F\|y-u\|(a)=u(b)+F|y(a)-u(a)|=u(b)+F|y(a)-A| \\ &\leqq u(b)+2F(C_1+2|y'(a)|)|a'-a| \\ &=u(b)+2F(C_1+2|u'(a)|)|a'-a|. \end{aligned} \tag{38}$$

If $|y'(a)|\leqq H(\varepsilon)$, then $y(b)\leqq\bar{B}(b)+\tfrac{1}{16}F^{-1}(C_1+2H(\varepsilon))\delta(\varepsilon)$. If $|y'(a)|>H(\varepsilon)$, then (35) holds. If (35i) holds, then

$$\begin{aligned} y(b)&\leqq\bar{B}(b)-\frac{1}{2}F^{-1}|u'(a)|+2F(C_1+2|u'(a)|)|a'-a| \\ &\leqq\bar{B}(b)-\frac{1}{2}F^{-1}|u'(a)|+\frac{1}{16}F^{-1}(C_1+2|u'(a)|)\delta(\varepsilon) \\ &=\bar{B}(b)+\frac{C_1}{16F}\delta(\varepsilon)-\frac{1}{8}F^{-1}(4-\delta(\varepsilon))|u'(a)|\leqq\bar{B}(b)+\frac{C_1}{16F}\delta(\varepsilon). \end{aligned}$$

On the other hand, if $|y'(a)|>H(\varepsilon)$ and (35ii) holds, then

$$|u'(b)| \geqq \frac{1}{2}F^{-1}|u'(a)|-|u_0'(b)| \geqq 2\frac{\max(\bar{B}(b-\delta(\varepsilon)), \bar{B}(b+\delta(\varepsilon)))+C_2\delta(\varepsilon)}{\delta(\varepsilon)},$$

and (34) holds no matter what the sign of $u'(b)$. Thus, using (34) in (38), we obtain

$$\begin{aligned} y(b) &\leqq \frac{\max(\bar{B}(b-\delta(\varepsilon)), \bar{B}(b+\delta(\varepsilon)))+C_2\delta(\varepsilon)-\frac{1}{2}|u'(b)|\delta(\varepsilon)}{1+\varepsilon} \\ &\quad +2F(C_1+2|u'(a)|)|a'-a| \\ &\leqq \frac{\max(\bar{B}(b-\delta(\varepsilon)), \bar{B}(b+\delta(\varepsilon)))+C_2\delta(\varepsilon)}{1+\varepsilon}+\frac{|u_0'(b)|\delta(\varepsilon)}{2(1+\varepsilon)} \\ &\quad -\frac{1}{4(1+\varepsilon)F}|u'(a)|\delta(\varepsilon)+\frac{F^{-1}}{16}(C_1+2|u'(a)|)\delta(\varepsilon) \\ &\leqq \frac{\max(\bar{B}(b-\delta(\varepsilon)), \bar{B}(b+\delta(\varepsilon)))}{1+\varepsilon}+\left(C_2+|u_0'(b)|+\frac{1}{16}C_1F^{-1}\right)\delta(\varepsilon) \\ &\quad -\frac{F^{-1}}{4(1+\varepsilon)}\left(1-\frac{1+\varepsilon}{2}\right)|u'(a)|\delta(\varepsilon) \\ &\leqq \frac{\max(\bar{B}(b-\delta(\varepsilon)), \bar{B}(b+\delta(\varepsilon)))}{1+\varepsilon}+\left(C_2+|u_0'(b)|+\frac{1}{16}C_1F^{-1}\right)\delta(\varepsilon). \end{aligned}$$

Thus if $|a'-a| \leqq \delta(\varepsilon)/(32F^2)$, then $y(b)$ is bounded above for y an arbitrary solution of (5) with $y(a')=A$.

Now consider the interval $[b-\delta_0/2, b+\delta_0/2]$ where δ_0 is as originally chosen. For any point in this interval the preceding procedure could be applied with $\delta_0/2$ replacing δ_0. In the procedure for any such point Δ, $\delta_0/2$, ε, and the introduced quantities, E, F, D, $\delta(\varepsilon)$, C_1, $G(\varepsilon)$, and C_2, would not change, and specifically $\delta(\varepsilon)/(32F^2C_2)$ would not change. Thus there is an interval, J, about a such that for $a' \in J$ and $b' \in [b-\delta_0/2, b+\delta_0/2]$ upper existence fails from a' to b'. □

If U (or L) is the set of points in the plane such that upper (or lower) existence fails from the first coordinate to every point in some neighborhood of the second coordinate, then Theorem 6 says that U (or L) is open. Theorem 7 will show that the bounding function, $\bar{B}(b'; a', A') \cdot (\underline{B}(b'; a', A'))$ is continuous on $U \times$ Reals ($L \times$ Reals).

THEOREM 7. *$\bar{B}(b'; a', A')(\underline{B}(b'; a', A'))$ is continuous for $(a', b') \in U$ (or L) and A' real.*

Proof. Consider the case of U, the other being similar. Let (a, b) in U, A real, and ε in $(0, 1)$ be given. Let J and I be closed bounded intervals about a and b respectively such that $J \times I \subset U$. First consider A' fixed at A.

Take Δ, an interval, and δ_0, a positive real, such that $[b-2\delta_0, b+2\delta_0] \subset I$, the $(2\delta_0)$-neighborhoods of J and I are contained in Δ, and first and second boundary value problems have unique solutions on subintervals of I of length at most $2\delta_0$. Take the introduced quantities of Theorem 6, except for $H(\varepsilon)$ and including $u_0(b'; a, A)$, with the additional restriction on $\delta(\varepsilon)$ that $|\bar{B}(b_2'; a, A)-\bar{B}(b_1'; a, A)| \leqq \varepsilon$, $|u_0(b_2'; a, A)-u_0(b_1'; a, A)| \leqq \varepsilon$, and $|u_0'(b_2'; a, A)-u_0'(b_1'; a, A)| \leqq \varepsilon$ whenever $|b_2'-b_1'| \leqq \delta(\varepsilon)$.

Then $\bar{B}(b'\pm\delta(\varepsilon); a, A)\leqq\bar{B}(b'; a, A)+\varepsilon\leqq\bar{B}(b; a, A)+2\varepsilon$ and $|u_0(b'; a, A)|\leqq|u_0(b; a, A)|+\varepsilon$ whenever $|b'-b|\leqq\delta(\varepsilon)$.

Let $H(\varepsilon, b')$ be as in (37) of Theorem 6 with b' replacing b and

$$H^*(\varepsilon)=\sup_{b'\varepsilon[b-\delta(\varepsilon),b+\delta(\varepsilon)]} H(\varepsilon, b')\leqq H(\varepsilon, b)+2F\max\left(2\varepsilon, \varepsilon+\frac{4\varepsilon}{\delta(\varepsilon)}\right)$$

$$\leqq H(\varepsilon, b)+\frac{5\varepsilon}{\delta(\varepsilon)}2F.$$

Then if $|a'-a|\leqq\delta(\varepsilon)/(16F^2[C_1+2(H^*(\varepsilon)+1)])$ and $|b'-b|\leqq\delta(\varepsilon)$ one has, as in the proof of Theorem 6,

$$|y'(a; a', A)|\leqq H(\varepsilon, b')\rightarrow y(b'; a', A)\leqq\bar{B}(b; a, A)+\varepsilon+\frac{1}{8}F^{-1}\delta(\varepsilon)$$

$$|y'(a; a', A)|>H(\varepsilon, b')\rightarrow y(b'; a', A)\leqq\max\left[\bar{B}(b; a, A)+\varepsilon+\frac{C_1}{16F}\delta(\varepsilon),\right.$$

$$\left.\frac{\bar{B}(b; a, A)+2\varepsilon}{1+\varepsilon}+\left(C_2+|u_0'(b; a, A)|+\varepsilon+\frac{1}{16}C_1F^{-1}\right)\delta(\varepsilon)\right]$$

where

$$\delta(\varepsilon)\leqq\frac{\varepsilon}{1+\varepsilon}[K\delta_0+L+k]^{-1}.$$

Thus there exists Q such that $|a'-a|\leqq\delta(\varepsilon)/(16F^2[C_1+2(H^*(\varepsilon)+1)])$ and $|b'-b|\leqq\delta(\varepsilon)$ assure $y(b'; a', A)\leqq\bar{B}(b; a, A)+Q\varepsilon$.

Now if in addition to the restriction on a' and b' one has $|A'-A|\leqq\varepsilon$ and the same derivative at a' for both solutions, then

$$y(b'; a', A')\leqq y(b'; a', A)+e^{k|b'-a'|}|A'-A|\leqq y(b'; a', A)+F|A'-A|$$

$$\leqq\bar{B}(b; a; A)+(Q+F)\varepsilon.$$

Thus $|a'-a|\leqq\delta(\varepsilon)/(16F^2[C_1+C_2(H^*(\varepsilon)+1)])$, $|b'-b|\leqq\delta(\varepsilon)$, and $|A'-A|\leqq\varepsilon$ assure that $\bar{B}(b'; a', A')\leqq\bar{B}(b; a, A)+(Q+F)\varepsilon$.

To complete the proof of continuity, again let A' be fixed at A and let ε be positive. Then there is a solution $y(x; a, A, m)$ such that $y(b; a, A, m)=\bar{B}(b; a, A)$. Now take the solution, $y(x; a', A, y'(a'; a, A, m))$, which is continuous in (x, a') since $y'(a'; a, A, m)$ is continuous in a'. Thus there exists positive δ such that $|y(b'; a', A, y'(a'; a, A, m))-y(b; a, A, y'(a; a, A, m))|\leqq\varepsilon$ whenever $|a'-a|+|b'-b|\leqq\delta$. However, $y(b; a, A, y'(a; a, A, m))=y(b; a, A, m)=\bar{B}(b; a, A)$. Thus $|a'-a|+|b'-b|\leqq\delta$ gives

$$y(b'; a', A, y'(a'; a, A, m))\geqq\bar{B}(b; a, A)-\varepsilon,$$

so

$$\bar{B}(b'; a', A)\geqq\bar{B}(b, a, A)-\varepsilon.$$

The perturbation from A to A' is treated as above. □

THEOREM 8. *If upper (lower) existence fails from a to every point b in an open set containing the compact set, C, then there is an interval J about a such that upper (lower) existence fails from a′ to b′ whenever a′ ∈ J and b′ ∈ C. The bounding function for*

(a', A') converges uniformly on C to the bounding function for (a, A) as (a', A') converges to (a, A).

Proof. Theorem 6 provides intervals $J_{b'}$ and $I_{b'}$ about a and b' respectively for every b' in C. Compactness of C yields a single interval J about a which works for C. For $a' \in J$, (a', A') in a closed bounded rectangle about (a, A) and b' in C, Theorem 7 gives uniform continuity of the bounding function in (a', A', b') and the second statement of the theorem follows. □

The functions, $\bar{B}_R(b'; a', A')$ and $\bar{B}_L(b'; a', A')$, are of course defined on U (or L), but they need not be continuous there since they may not be continuous in b'. However, partial analogues of Theorem 7 and 8 may be obtained.

THEOREM 9. *Let a, b, A, ε be given such that $(a, b) \in U$ (or L) and $\varepsilon > 0$. Then there exists positive δ such that $|a'-a|$, $|b'-b|$, $|A'-A|$ all less than δ assures $\bar{B}_L(b; a, A) - \varepsilon \leqq \bar{B}_L(b'; a', A') \leqq \bar{B}_R(b'; a', A') \leqq \bar{B}_R(b; a, A) + \varepsilon$ $[\underline{B}_R(b; a, A) - \varepsilon \leqq \underline{B}_R(b'; a', A') \leqq \underline{B}_L(b'; a', A') \leqq \underline{B}_L(b; a, A) + \varepsilon]$. If $\bar{B}_R(b; a, A) = \bar{B}_L(b; a, A)$ $[\underline{B}_R(b; a, A) = \underline{B}_L(b; a, A)]$, that is, (a, b, A) is a point of continuity with respect to the second position for the two derivatives, then (a, b, A) is a point of continuity for the two derivatives.*

Proof. The case for U is treated. Let a, b, A, ε be given with $(a, b) \in U$ and $\varepsilon > 0$.

If the first statement fails for $\bar{B}_R(b'; a', A')$, then there exists a sequence $\{a'_n, b'_n, A'_n\}$ converging to (a, b, A) for which the required condition on $\bar{B}_R(b'_n; a'_n, A'_n)$ is violated.

Let y_n be a solution of (5) with $y_n(a'_n) = A'_n$, $y_n(b'_n) = \bar{B}(b'_n; a'_n, A'_n)$, and $y'_n(b'_n) = \bar{B}_R(b'_n; a'_n, A'_n)$. Since $\bar{B}(b'; a', A')$ is continuous in (a', b', A'), the sequence $\{\bar{B}(b'_n; a'_n, A'_n)\}$ converges to $\bar{B}(b; a, A)$.

Applying Lemma 2 to y_n with a, x replaced by b'_n, $b'_n + h$, one has

$$|y_n(b'_n+h) - y_n(b'_n) - y'_n(b'_n)h| \leqq [D + K|y_n(b'_n)| + (K|h| + L)|y'_n(b'_n)|]k^{-1}(e^{k|h|} - 1)|h|$$

where D_x of (17) has been replaced by some common bound valid for a suitably large interval. Then, on taking $|h|$ small enough and n large enough that $(a'_n, b'_n + h) \in U$ and $(a, b+h) \in U$, we obtain

$$\bar{B}(b'_n+h; a'_n, A'_n) \geqq y_n(b'_n+h) \geqq [1 - (K|h| + L)k^{-1}(e^{k|h|} - 1)\,\mathrm{sign}\,(y'_n(b'_n)h)]y'_n(b'_n)h$$
$$-[1 + Kk^{-1}(e^{k|h|} - 1)|h|]|y_n(b'_n)| - Dk^{-1}(e^{k|h|} - 1)|h|.$$

If the $y'_n(b'_n)$'s are unbounded above, take h positive and sufficiently small that the coefficient of $y'_n(b'_n)h$ is at least $\frac{1}{2}$. Then the $\bar{B}(b'_n+h; a'_n, A'_n)$'s are unbounded above since the sequence $\{y_n(b'_n)\}$ converges to $\bar{B}(b; a, A)$. However, $\{\bar{B}(b'_n+h; a'_n, A'_n)\}$ converges to $\bar{B}(b+h; a, A)$, a contradiction. Similarly the $y'_n(b'_n)$'s must be bounded below (taking h negative), and by extracting a subsequence one can consider $\{y'_n(b'_n)\}$ to be convergent.

Now $\{y_n\}$ and $\{y'_n\}$ must converge to y and y' where y is a solution of (5) and (14) with $y(b) = \bar{B}(b; a, A)$, the convergence being uniform on any finite interval. Then $\{y'_n(b'_n)\}$ must converge to $y'(b)$ which is in $[\bar{B}_L(b; a, A), \bar{B}_R(b; a, A)]$. This contradicts $\bar{B}_R(b'_n; a'_n, A'_n)$ violating the required condition for all n, and the original assumption must be false. Thus the first statement holds for $\bar{B}_R(b'; a', A')$. The proof for $\bar{B}_L(b'; a', A')$ is similar.

The second statement follows immediately from the first, with the parenthetical remark provided by Theorem 5. □

THEOREM 10. *If in addition to the hypotheses of Theorem* 8, *the bounding function for* (a, A) *has a derivative at every point of C, then as* (a', A') *converges to* (a, A) *both right and left derivatives of the bounding function for* (a', A') *converge uniformly on C to the derivative of the bounding function for* (a, A).

Proof. With J the interval of Theorem 8, restrict a' to J, (a', A') to a closed bounded rectangle about (a, A), and b' to C. Then given $\varepsilon > 0$ and b in C, by Theorem 9 there exists a positive δ_b such that $|a'-a|$, $|A'-A|$, and $|b'-b|$ all less than δ_b guarantees $|\bar{B}_R(b'; a', A') - \bar{B}_R(b; a, A)| < \varepsilon/2$. A finite number of the intervals of type $|b'-b| < \delta_b$ cover C, say for $b_1, b_2, \cdots, b_n$. Let δ be positive but smaller than each δ_{b_i}.

If now $|a'-a|$ and $|A'-A|$ are less than δ, for b' in C there exists b_i such that $|b'-b_i| < \delta_{b_i}$. Then $\bar{B}_R(b'; a', A')$ and $\bar{B}_R(b'; a, A)$ differ from $\bar{B}_R(b_i; a, A)$ by at most $\varepsilon/2$, and from each other by at most ε. □

THEOREM 11. *Given a, b, and A, then there exists an open interval I about b such that upper existence fails at every b' in I or lower existence fails at every b' in I if and only if there exists a B and a neighborhood,* I^*, *of b such that no solution of* (5) *through* (a, A) *crosses* $I^* \times \{B\}$.

Proof. Since with either type of existence failure on I the bounding function corresponding to (a, A) is continuous on I it is clear that the second condition follows from the first.

Suppose the second condition holds. Then every solution of (5) through (a, A) must remain below B on I^* or must remain above B on I^*. Since two solutions with one remaining above and one remaining below would lead to a solution through (b, B), either all must remain above or all must remain below. Thus for the family of solutions, either upper existence must fail on I^* or lower existence must fail on I^*. □

THEOREM 12. *If there is some neighborhood of* (b, B) *which is not entered by any solution of* (5) *through* (a, A), *then there is some neighborhood, N, of* (a, A) *such that no solution of* (5) *through* (b, B) *satisfies a boundary condition corresponding to a point of N.*

Proof. By Theorem 11 there is an interval I about b such that, for the family of solutions corresponding to a and A, either upper existence fails at every point of I or lower existence fails at every point of I. Also (b, B) is on the unattainable side of the bounding curve. Theorem 6 gives the bounding function defined in some neighborhood of (a, b, A), and the continuity of the bounding function from Theorem 7, assures that for (a', A') in a sufficiently small neighborhood of (a, A) the value of the bounding function at (a', b, A') will remain on the proper side of B. Thus no solution of (5) through an (a', A') in the indicated neighborhood of (a, A) passes through (b, B). □

THEOREM 13. *If upper (lower) existence fails from a to b' for every b' in some interval about b, then there exists an interval, J, about a such that either upper existence fails from b to every point a' in J or lower existence fails from b to every point a' in J.*

Proof. If $B > \bar{B}(b; a, A)$, then (b, B) satisfies the hypotheses of Theorem 12. Thus b, a, and B (for a, b, and A) satisfy the second condition of Theorem 11, with A replacing B. The equivalent first condition of Theorem 11 is the desired conclusion. □

It should be noted that if the boundary condition, $y'(b) = n$, is considered, then the results corresponding to Theorem 6–10, 12, 13 may not hold. In Example 1, $C(x)$ may vanish on an interval J and if b^* is any point of the interval then a is a point of existence failure for the problem with y' specified as 0 at b^* and a value for y to be attained at a. Since the differential equation is actually linear, a must be an isolated point of existence failure. Thus if a' is sufficiently close to a then b^* is an existence point for a' for the original problem. Thus U (or L) does not contain a neighborhood of (a, b^*); and if upper (lower) bounding functions are thought of as having value $\infty(-\infty)$ where upper (lower) existence holds, then continuity and convergence properties fail at (a, b^*).

Indeed, since the general solution in Example 1 satisfies $y'(x) = C(x)y(x) + Q$ for $x \geqq b$, Q a constant, if Z is any compact subset of the zeros of $C(x)$, then for a' sufficiently close but not equal to a existence will hold from a' to every point in some open set about Z.

Theorem 11 will hold in the converse direction, but the first condition will imply the second only if every bounding function on intervals is locally bounded. A counter example would then have to be sought amonst nonlinear equations.

Acknowledgment. The authors wish to thank one of the referees for significant suggestions concerning earlier versions of this paper.

REFERENCES

[1] E. A. CODDINGTON AND N. LEVINSON, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[2] P. BAILEY, L. SHAMPINE AND P. WALTMAN, *Nonlinear Two Point Boundary Value Problems*, Academic Press, New York, 1968.

[3] E. W. HOBSON, *The Theory of Functions of a Real Variable*, Dover, New York, 1957.

SIAM J. MATH. ANAL.
Vol. 9, No. 3, June 1978

Copyright © 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0903-0011$1.00/0

ON THE SPECTRAL RADIUS OF INTEGRAL OPERATORS DEFINED BY A CLASS OF DIFFERENCE KERNELS*

H. D. VICTORY, JR.†

Abstract. We consider the problem of estimating the spectral radius of integral operators defined by a certain class of difference kernels. Under the conditions that the kernels are summable, positive, even about $x = 0$, and decreasing in $|x|$, we obtain sharp bounds for the spectral radius of the associated integral operator by using the Perron–Frobenius–Jentsch characterization of the dominant eigenvalue. Moreover, we show that our bounds determine asymptotics for the spectral radius, which agree with the results of H. Widom, who used the underlying self-adjointness for his results. We show how techniques in this paper work for systems of such integral equations which lead to problems where self-adjointness may be absent.

1. Introduction. Many problems in applied mathematics lead to a study of inhomogeneous integral equations of the following type:

$$f(x) = \gamma \int_0^\tau k(x-y)f(y)\,dy + g(x), \tag{1.1}$$

for $0 \leqq \tau \leqq \infty$, $0 \leqq \gamma < \infty$. The kernel function will be assumed to be positive, summable, decreasing in $|x|$, and even about $x = 0$. Kernels with these properties appear in the theories of radiative transfer, neutron transport, gas dynamics (Couette flow), electrodynamic wave diffraction [11] and often have the Laplace transform representation

$$k(x-y) = \int_\alpha^\infty \psi(t)\, e^{-|x-y|t} \frac{dt}{t},$$

with α, $\psi(t) \geqq 0$, and ψ satisfying a Hölder condition on (α, ∞). Also such difference kernels arise in the work of Hardy, Littlewood, and Pólya [7] when the assumption of symmetry is added to their hypotheses [7, p. 227, Thm. 318].

We write (1.1) in abbreviated notation as

$$f(x) = \gamma \Lambda_\tau f(x) + g(x). \tag{1.2}$$

Under the hypotheses on $k(x)$, it can be shown that Λ_τ is a bounded linear operator on $L^p(0, \tau)$, $p \geqq 1$, and on $C[0, \tau]$ where

$$\begin{aligned} L^p(0,\tau) &= \left\{ f \colon \int_0^\tau |f(x)|^p\,dx < \infty \right\}, \\ C[0,\tau] &= \{ f \colon f \text{ continuous on } [0,\tau] \}. \end{aligned} \tag{1.3}$$

The first result is derived by use of the fact that the convolution with an L^1 function is a bounded linear mapping of L^p into itself. The same result for $C[0, \tau]$ follows trivially from the summability of $k(x)$. For $\tau < \infty$, Λ_τ will be a compact linear operator on $L^p(0, \tau)$, since k can be approximated in the L^1 norm by C^∞ functions with compact support. The operators generated by these will converge to Λ_τ in operator norm and are compact. Moreover, due to the positivity of $k(x)$, we can see that Λ_τ leaves invariant the cone of nonnegative functions in $L^p(0, \tau)$ and $C[0, \tau]$.

This paper is written with two objectives in mind. One is to use the Perron–Frobenius–Jentsch characterization of the dominant eigenvalue [8, p. 929] to obtain

* Received by the editors May 21, 1976, and in final revised form October 10, 1976.

† Texas Tech University, Lubbock, Texas 79409. This research was supported by National Science Foundation under Grant Eng. 75-08407.

sharp upper and lower bounds for it. As is well known, [8, p. 924], the dominant eigenvalue is equal to the spectral radius of Λ_τ defined as

$$\|\Lambda_\tau\|_{sp} = \lim_{n\to\infty} \|\Lambda_\tau^n\|^{1/n}. \tag{1.4}$$

From these upper and lower bounds, we can obtain asymptotics for $\|\Lambda_\tau\|_{sp}$ when studied as a perturbation of $\|\Lambda_\infty\|_{sp}$. The other objective is to show how this technique can be applied to systems of integral operators where $k(x)$ is an $N \times N$ matrix, each component of which satisfies the assumptions discussed above.

For the scalar case, Λ_τ is a compact self-adjoint operator when defined on $L^2(0, \tau)$. H. Widom [15, pp. 401–402, 412] made crucial use of these properties in obtaining asymptotics for the positive eigenvalues of Λ_τ. In our work here, we do not exploit the self-adjointness of Λ_τ in the scalar case. However, our techniques work for systems of equations, where Widom's methods do not apply, since self-adjointness is rarely present.

We remark that the Perron–Frobenius–Jentsch theory has been exploited to advantage by H. S. Wilf and N. G. de Bruijn [16, pp. 32–34] in their analysis of the Hilbert kernel. Indeed this paper shows that their techniques are easily generalized to a wider class of kernels. Moreover, in § 3, we show that the class of kernels considered enables us to further analyze the parameter in p in $f_2(x) = \cos \pi x/(\tau + p)$, used in [16] to obtain upper bounds.

In § 2, we discuss properties of $\|\Lambda_\tau\|_{sp}$ and its associated eigenfunction. We show we can get sharp lower bounds on $\|\Lambda_\tau\|_{sp}$ when k fulfills the assumptions discussed above. In § 3, we show that we can sharpen the upper bound on $\|\Lambda_\tau\|_{sp}$ with a few more assumptions on $k(x)$. In § 4, we generalize our previous results to systems of integral equations, and give applications. An analysis of some classical kernels is given in § 5.

We would like to mention that this work was inspired by the results of J. Bolmarcich [1].

2. General properties of $\|\Lambda_\tau\|_{sp}$. In the Introduction, we have remarked that Λ_τ leaves the cone of nonnegative continuous functions invariant. We can say more: The positivity of $k(x)$ implies Λ_τ will map a nonnegative function to a positive one. We will describe Λ_τ as a *strictly positive* oerator, following Karlin [8, p. 920]. These observations lead to the following:

THEOREM 2.1. *The maximum eigenvalue of Λ_τ is equal to its spectral radius, and its eigenfunction ϕ_τ is of simple multiplicity, positive, and even about $\tau/2$.*

Proof. We have seen that Λ_τ is a compact linear operator on $C[0, \tau]$. Because the cone K of nonnegative continuous functions has an interior, and Λ_τ maps K into its interior from the remark above, we can deduce that its largest eigenvalue is real, of simple multiplicity, and equal to $\|\Lambda_\tau\|_{sp}$ [8, p. 924]. The eigenfunction is positive for $0 \leqq x \leqq \tau$.

A straightforward calculation, using the evenness of k about $x = 0$, shows that ϕ_τ is even about $\tau/2$. This completes the proof of the theorem.

The next result gives the behavior of $\|\Lambda_\tau\|_{sp}$ as a function of τ. For $\tau = \infty$, $\|\Lambda_\infty\|_{sp} = \int_{-\infty}^{\infty} k(x)\, dx$, a fact which follows from norm estimates on $\Lambda_\infty f$ and from the results of M. G. Krein [9, p. 224] which state that the spectrum of Λ_∞ includes

$$\{\lambda : \lambda = \hat{k}(\xi), \xi \text{ real}\},$$

where $\hat{k}$ is the Fourier transform of k, defined in (2.2).

We will consider Λ_τ defined on $C[0, \tau]$. The proof of the following theorem exploits the Perron–Frobenius–Jentsch characterization of $\|\Lambda_\tau\|_{sp}$ [8, p. 929],

$$(2.1)\qquad \begin{aligned} \|\Lambda_\tau\|_{sp} &= \sup\{\lambda > 0: \exists f \geqq 0: \Lambda_\tau f \geqq \lambda f\}, \\ \|\Lambda_\tau\|_{sp} &= \inf\{\lambda > 0: \exists f \geqq 0: \Lambda_\tau f \leqq \lambda f\}, \end{aligned}$$

and is similar to that proved by T. W. Mullikin [12, p. 513–514] for a class of strictly positive operators including those analyzed in this paper. We refer the reader to [12] for the main ideas in the proof of the following:

THEOREM 2.2. $\|\Lambda_\tau\|_{sp}$ *is a strictly monotone increasing, continuous function of* τ.

The next result will produce a lower bound on $\|\Lambda_\tau\|_{sp}$ which will imply that

$$\lim_{\tau\to\infty} \|\Lambda_\tau\|_{sp} = \|\Lambda_\infty\|_{sp}.$$

The techniques used in the proof of the following are similar to those used by Wilf and de Bruijn [16, p. 32–34] in their analysis of the Hilbert kernel. To generalize their results, one merely exploits the decreasing nature of $k(x)$ and the fact that, for even functions, the Fourier transform is the same as the Fourier cosine transform.

THEOREM 2.3. $\|\Lambda_\tau\|_{sp} \geqq \hat{k}(\pi/\tau)$ *where* $\hat{k}$ *denotes the Fourier transform of* k, *defined as*

$$(2.2)\qquad \hat{k}(\xi) = \int_{-\infty}^{\infty} e^{ix\xi} k(x)\, dx.$$

Proof. We use (2.1) along with the choice of $f_1 = \cos \pi x/\tau$ to produce the lower bound for $\|\Lambda_\tau\|_{sp}$. We refer the reader to [16] for details.

3. Upper bounds for $\|\Lambda_\tau\|_{sp}$. With a few more assumptions on $k(x)$, we can produce sharp upper bounds on $\|\Lambda_\tau\|_{sp}$. We have our main result:

THEOREM 3.1. *Let* $k(x)$ *be positive, even about* $x = 0$, *decreasing in* $|x|$, *and summable. If* $k(x) = d^2k_3/dx^2$, *where* k_3 *is decreasing and positive with* $\lim_{x\to\infty} k_3(x) = 0$, *and such that*

$$(3.1)\qquad \mu^* = \sup_{\mu} \left\{ \frac{d}{dx}(e^{\mu x} k_3(x)) \leqq 0, x \geqq 0 \right\} > 0,$$

then

$$(3.2)\qquad \|\Lambda_\tau\|_{sp} \leqq \hat{k}(\pi/\tau + p(\tau)),$$

where p *is a continuous function of* τ *for* τ *sufficiently large with*

$$2/\mu^* \leqq p(\tau) \leqq \pi/\mu^*.$$

Proof. We again use (2.1) with $f_2 = \cos \pi x/(\tau + p)$ to give our upper bound. Now

$$(3.3)\qquad \begin{aligned} \int_{-\tau/2}^{\tau/2} k(x-y) \cos\frac{\pi y}{\tau+p}\, dy = \int_{-\infty}^{\infty} k(x-y) \cos\frac{\pi y}{\tau+p}\, dy \\ - \int_{\tau/2}^{\infty} k(x-y) \cos\frac{\pi y}{\tau+p}\, dy \\ - \int_{-\infty}^{-\tau/2} k(x-y) \cos\frac{\pi y}{\tau+p}\, dy. \end{aligned}$$

We claim that

$$(3.4)\qquad \int_{\tau/2}^{\infty} k(x-y)\cos\frac{\pi y}{\tau+p}\,dy \geqq 0 \quad\text{and}\quad \int_{-\infty}^{-\tau/2} k(x-y)\cos\frac{\pi y}{\tau+p}\,dy \geqq 0.$$

We prove only the first inequality; the proof for the second is identical. By an argument similar to that used in Theorem 2.3 [see 16, p. 34] we can conclude that

$$\int_{3(\tau+p)/2}^{\infty} k(x-y)\cos\frac{\pi y}{\tau+p}\,dy > 0.$$

So

$$(3.5)\qquad \int_{\tau/2}^{\infty} k(x-y)\cos\frac{\pi y}{\tau+p}\,dy > \int_{\tau/2}^{3(\tau+p)/2} k(x-y)\cos\frac{\pi y}{\tau+p}\,dy.$$

An easy calculation yields

$$(3.6)\qquad \begin{aligned}\int_{\tau/2}^{3(\tau+p)/2} k(x-y)\cos\frac{\pi y}{\tau+p}\,dy = \frac{1}{(\tau+p)}\Big[&-k_2(\tau/2-x)\cos\frac{\pi\tau}{2(\tau+p)}(\tau+p)\\ &-\pi k_3(\tfrac{3}{2}(\tau+p)-x)-\pi k_3(\tau/2-x)\sin\frac{\pi\tau}{2(\tau+p)}\\ &-\frac{\pi^2}{(\tau+p)}\int_{\tau/2}^{3(\tau+p)/2} k_3(y-x)\cos\frac{\pi y}{\tau+p}\,dy\Big],\end{aligned}$$

where

$$k_2(x) = -\int_x^{\infty} k(t)\,dt.$$

We wish to estimate the latter integral in (3.6). For $y\in[\tau/2, (\tau+p)/2]$, $\cos \pi y/(\tau+p) \geqq 0$, and we get

$$(3.7)\qquad \begin{aligned}&\frac{\pi^2}{(\tau+p)}\left\{\int_{\tau/2}^{(\tau+p)/2} k_3(y-x)\cos\frac{\pi y}{\tau+p}\,dy + \int_{\tau+p/2}^{3(\tau+p)/2} k_3(y-x)\cos\frac{\pi y}{(\tau+p)}\,dy\right\}\\ &\quad\leqq \frac{\pi^2}{\tau+p}\left\{k_3(\tau/2-x)\frac{(\tau+p)}{\pi}\sin\frac{\pi y}{\tau+p}\Big|_{\tau/2}^{(\tau+p)/2}\right.\\ &\qquad\left. + k_3(\tfrac{3}{2}(\tau+p)-x)\frac{\tau+p}{\pi}\sin\frac{\pi y}{\tau+p}\Big|_{(\tau+p)/2}^{3(\tau+p)/2}\right\}.\end{aligned}$$

So

$$(3.8)\qquad \begin{aligned}&-\frac{\pi^2}{\tau+p}\int_{\tau/2}^{3(\tau+p)/2} k_3(y-x)\cos\frac{\pi y}{\tau+p}\,dy\\ &\quad\geqq -\pi k_3(\tau/2-x)+\pi k_3(\tau/2-x)\sin\frac{\pi\tau}{2(\tau+p)}+2\pi k_3(\tfrac{3}{2}(\tau+p)-x).\end{aligned}$$

With (3.8), we have that

$$(3.9)\qquad \begin{aligned}\int_{\tau/2}^{3(\tau+p)/2} k(y-x)\cos\frac{\pi y}{\tau+p}\,dy \geqq \frac{1}{\tau+p}\Big\{&-k_2(\tau/2-x)(\tau+p)\cos\frac{\pi\tau}{2(\tau+p)}\\ &-\pi k_3(\tau/2-x)+\pi k_3(\tfrac{3}{2}(\tau+p)-x)\Big\}.\end{aligned}$$

To insure this positive, we must demand that

$$-k_2(\tau/2-x)(\tau+p)\cos\frac{\pi\tau}{2(\tau+p)}-\pi k_3(\tau/2-x)\geqq 0 \tag{3.10}$$

for $|x|\leqq\tau/2$. This will be true if we can find p as a function of τ for which

$$\frac{d}{dx}\left(\exp\left(\frac{\pi}{\tau+p}\csc\frac{\pi p}{2(\tau+p)}x\right)k_3(x)\right)\leqq 0 \quad \text{for } x\geqq 0. \tag{3.11}$$

From (3.1), we see that we must be able to find $p=p(\tau)$ such that

$$\frac{\pi}{\tau+p}\csc\frac{\pi p}{2(\tau+p)}=\mu^*, \quad \text{or} \quad \frac{(\tau+p)}{\pi}\sin\frac{\pi p}{2(\tau+p)}=\frac{1}{\mu^*} \tag{3.12}$$

Let us define $\eta=1/\tau$. Then, as a function of η, p, we have

$$F(\eta,p)=\frac{1+\eta p}{\eta\pi}\sin\frac{\pi\eta p}{2(1+\eta p)}-\frac{1}{\mu^*}=0. \tag{3.13}$$

For $\eta=0$, we have that $p=2/\mu^*$. We can easily see that F, F_η, F_p are continuous near $\eta=0$, $p=2/\mu^*$, and that

$$F_p(0,2/\mu^*)=\pi/2\neq 0.$$

From the implicit function theorem, we can solve for p as a continuous function of η, η suitably small, or as a continuous function of τ.

Professor G. M. Wing [17] has suggested the following argument to obtain upper and lower bounds on $p(\tau)$. We first show that we can define $p(\tau)$ for all τ: Suppose we have $p(\tau)$ a continuous function of τ for $\tau\in[\tau_0,\infty)$. Can we extend $p(\tau)$ to a continuous function of $[\tau_0-\varepsilon,\infty)$? If $F_p(\tau,p)\neq 0$ at (τ_0,p_0), where $F(\tau,p)$ is defined in (3.13), then such an extension can be made by the implicit function theorem. An easy calculation shows that

$$F_p(\tau_0,p_0)=\frac{1}{\pi}\sin\left(\frac{\pi p_0}{2(\tau_0+p_0)}\right)+\frac{\tau_0}{2(\tau_0+p_0)}\cos\frac{\pi p_0}{2(\tau_0+p_0)}. \tag{3.14}$$

This cannot be equal to zero, since otherwise,

$$\tan\frac{\pi p_0}{2(\tau_0+p_0)}=-\frac{\pi\tau_0}{2(\tau_0+p_0)},$$

and this contradicts the fact that $0\leqq\pi p_0/(2(\tau_0+p_0))<\pi/2$, since $p_0>0$. So $p(\tau)$ is defined and continuous for all $\tau\geqq 0$.

To get our bounds for $p(\tau)$, we define

$$\chi(\tau)=\frac{\pi}{2}\frac{p(\tau)}{\tau+p(\tau)}.$$

Equation (3.12) becomes

$$\frac{\sin\chi(\tau)}{\chi(\tau)}=\frac{2}{\mu^*p(\tau)}. \tag{3.15}$$

Because $[\sin\chi(\tau)]/(\chi(\tau))\leqq 1$, we get that $p(\tau)\geqq 2/\mu^*$, which produces our lower bound.

We get our upper bound by noting that $[\sin w]/(w)$ is a decreasing function of w for $0 \leqq w \leqq \pi/2$. From (3.15), we get

$$p(\tau) = 2/\mu^*\left(\frac{\sin \chi(\tau)}{\chi(\tau)}\right). \tag{3.16}$$

We observe that $x(0) = \pi/2$, and hence $[\sin \chi(\tau)]/(\chi(\tau)) \geqq 2/\pi$. Thus $p(\tau) \leqq \pi/\mu^*$.

We can draw the conclusion that

$$\int_{-\tau/2}^{\tau/2} k(x-y) \cos \frac{\pi y\, dy}{\tau + p(\tau)} \leqq \int_{-\infty}^{\infty} k(x-y) \cos \frac{\pi y}{\tau + p(\tau)}\, dy = \hat{k}\left(\frac{\pi}{\tau + p(\tau)}\right) \cos \frac{\pi x}{\tau + p(\tau)},$$

and the theorem follows from (2.1).

Remark. From (3.1), we see that the class of kernels being considered are those for which $\mu^* > 0$.

4. Generalization to systems. In general self-adjointness is absent in the study of $N \times N$ systems. We consider (1.1) where $k(x-y)$ is a matrix whose entries have the properties described in § 1. We will use the following Banach space of vector functions:

$$C_N[0, \tau] = \left\{ f: f_i \text{ continuous}, \|f\| = \max_i \max_{[0,\tau]} |f_i(x)| \right\}. \tag{4.1}$$

A partial ordering in this space is induced by the cone $\mathcal{K}_N$ of vector functions possessing nonnegative entries. An operator Λ is strictly positive if Λf is a vector function whose entries are positive on $[0, \tau]$.

Our main result will consider an operator Λ_R for which

$$\Lambda_R f(x) = \int_{-\infty}^{\infty} k(x-y) f(y)\, dy. \tag{4.2}$$

A power N_0 of Λ_R is assumed to be strictly positive. It is easy to deduce that $\Lambda_\tau^{N_0}$ is strictly positive when τ is sufficiently large.

If we denote the Fourier transform of k by $\hat{k}$, we can conclude that $\hat{k}(0)$ will be a nonnegative matrix, a power of which is strictly positive. From the assumptions on Λ_R, we have that $\Lambda_R^{N_0} k$ will have positve entries, each of which is summable. Also

$$\begin{aligned} 0 < \int_{-\infty}^{\infty} \Lambda_R^{N_0} k(x)\, dx &= \Lambda_R^{N_0} k(0) = k * \Lambda_R^{N_0 - 1}(0) \\ &= \hat{k}(0) \Lambda_R^{N_0 - 1} k(0) = (\hat{k}(0))^{N_0 + 1}. \end{aligned} \tag{4.3}$$

These equalities follow from the fact that iterates of Λ_R are convolutions and from properties of the Fourier transform of convolutions.

Because k has nonnegative and summable entries, $\hat{k}(\xi)$ will be a nonnegative matrix for ξ near zero, and $\hat{k}^{N_0+1}(\xi)$ will be strictly positive [8, p. 924]. Hence both $\|\hat{k}(0)\|_{sp}$ and $\|\hat{k}(\xi)\|_{sp}$ will be the largest eigenvalues of $\hat{k}(0)$ and $\hat{k}(\xi)$ respectively, and these are of simple multiplicity. The results of Karlin [8, pp. 930, 932] apply:

(i) $\|\hat{k}(\xi)\|_{sp} \to \|\hat{k}(0)\|_{sp}$.

(ii) The entries of $u_R(\xi)$ approach those of $u_R(0)$ where $u_R(\xi)$ and $u_R(0)$ are the positive eigenvectors corresponding to $\|\hat{k}(\xi)\|_{sp}$ and $\|\hat{k}(0)\|_{sp}$ respectively.

We have the following theorem:

THEOREM 4.1. *Let the matrix $k(x)$ have entries which satisfy the hypotheses in Theorem* 2.3 *and* 3.1 *with* (3.1) *replaced by*

$$\mu^* = \inf_{\substack{i=1,\cdots,N \\ j=1,\cdots,N}} \sup_{\mu_{ij}} \left\{ \frac{d}{dx}\left(e^{\mu_{ij}x} k_3^{ij}(x)\right) \leqq 0, x \geqq 0 \right\}. \tag{4.4}$$

(For those entries for which $k_{ij}(x) \equiv 0$, we take $\mu_{ij} = \infty$.) Then

(4.5)
- (i) $\|\Lambda_\tau\|_{\text{sp}}$ *is a continuous, monotone increasing function of* τ.
- (ii) $\|\Lambda_\tau\|_{\text{sp}}$ *is simple, and its associated eigenfunction is even about* $\tau/2$.
- (iii) $\|\hat{k}(\pi/\tau)\|_{\text{sp}} \leqq \|\Lambda_\tau\|_{\text{sp}} \leqq \|\hat{k}(\pi/(\tau + p(\tau)))\|_{\text{sp}}$,

where $p(\tau)$ is a bounded continuous function of τ.

The proofs are a trivial modification of those for the scalar case. For example, in showing $\|\hat{k}(\pi/\tau)\|_{\text{sp}} \leqq \|\Lambda_\tau\|_{\text{sp}}$, we would show that the signs of the components of

$$\int_{\tau/2}^{\infty} k(x-y) u_R\left(\frac{\pi}{\tau}\right) \cos\frac{\pi y}{\tau}\, dy$$

are negative for the upper bound; instead of (3.11), we would show the components of

$$\frac{d}{dx}\left(\exp\left(\frac{\pi}{\tau+p} \csc\frac{\pi p}{2(\pi+p)} x\right) k_3(x)\right) u_R\left(\frac{\pi}{\tau+p}\right)$$

are negative for the upper bound; instead of (3.11), we would show the components of analysis.

Applications: *Multigroup neutron transport theory.* The kernel

$$k(x) = \frac{1}{2}\int_1^\infty e^{-t|x|\Sigma} \frac{dt}{t} C, \tag{4.6}$$

arises in the analysis of isotropically scattering slabs when discrete energy levels are assumed [2]–[5], [10], [13]. Here C is an $N \times N$ matrix, nonnegative, with $C^{N_0} > 0$. Σ is a diagonal matrix whose entries are ordered in the following manner: $1 \leqq \sigma_{11} \leqq \sigma_{22} \leqq \cdots \leqq \sigma_{NN}$. It can be shown that $\Lambda_R^{N_0}$ is a strictly positive operator from the assumptions on C. The entries of $k(x)$ satisfy the hypotheses in Theorem 4.1.

We observe that

$$k_3(x) = \frac{1}{2}\int_1^\infty e^{-t\Sigma x} \frac{dt}{\sigma_{ii}^2 t^3} C. \tag{4.7}$$

We must select the value of μ^* for which the components of

$$e^{\mu x} k_3(x) = \frac{1}{2}\int_1^\infty e^{(\mu - \sigma_{ii} t)x} \frac{dt}{\sigma_{ii}^2 t^3} C \tag{4.8}$$

are decreasing in x. For each i, j we see that $\mu < \sigma_{ii}$ will insure this. From an examination of asymptotics for $k_3(x)$, similar to that in [6, p. 28],

$$k_3^{ij}(x) \sim \frac{1}{\sigma_{ii} x} e^{-\sigma_{ii} x} C_{ij}, \tag{4.9}$$

and we see $\mu=\sigma_{ii}$ is sharp. Hence $\mu^*=1$, and, moreover, $p(\tau)$ can be obtained as perturbation of $2/\mu^*$ with terms bounded in τ, for τ large.

A perturbation argument similar to that of Rellich in [14, pp. 60–61] shows that

$$\|k(x)\|_{sp}=\|k(0)\|_{sp}-\tfrac{1}{2}x^2\frac{|u_L(0)\hat{k}''(0)u_R(0)|}{u_L(0)u_R(0)}+o(x^2) \tag{4.10}$$

where $u_L(0)\hat{k}(0)=\|\hat{k}(0)\|_{sp}u_L(0)$, and $u_L(0)$ is considered a row vector. Hence

$$\|\Lambda_\tau\|_{sp}=\|\hat{k}(0)\|_{sp}-\frac{\pi^2}{2\tau^2}\frac{|u_L(0)\hat{k}''(0)u_R(0)|}{u_L(0)u_R(0)}+o(\tau^{-2}). \tag{4.11}$$

So our techniques enable us to estimate the spectral radius for a class of non self-adjoint operators which includes those with important physical applications. The techniques of [15] cannot be employed here.

5. Analysis of some classical kernels. In this section, we will consider three integral operators which are self-adjoint when considered on $L^2(0,\tau)$. It is interesting to see how the Perron–Frobenius–Jentsch theory can be used to provide upper and lower bounds for $\|\Lambda_\tau\|_{sp}$ in lieu of variational techniques.

I. *Radiative transfer* [6]. We have

$$k(x)=\int_1^\infty \psi(t)\,e^{-|x|t}\frac{dt}{t}, \tag{5.1}$$

and hence

$$k_3(x)=\int_1^\infty \psi(t)\,e^{-|x|t}\frac{dt}{t^3}, \tag{5.2}$$

where $\psi(t)$ has the properties described in the Introduction. It is easy to see that $k(x)$ satisfies the hypotheses of Theorem 2.3 and Theorem 3.1. We must select the highest value of μ for which

$$\frac{d}{dx}(e^{\mu x}k_3(x))=\frac{d}{dx}\int_1^\infty \psi(t)\,e^{x(\mu-t)}\frac{dt}{t^3}\leqq 0, \qquad x\geqq 0. \tag{5.3}$$

An examination of the analysis of the multigroup transport kernel shows $\mu^*=1$. So $p(\tau)$ can be studied as a perturbation of $p(\infty)=2$, and we have

$$\hat{k}(\pi/\tau)\leqq\|\Lambda_\tau\|_{sp}\leqq\hat{k}\left(\frac{\pi}{\tau+p(\tau)}\right). \tag{5.4}$$

Noting that $p(\tau)$ is uniformly bounded in τ, we get for asymptotics for $\|\Lambda_\tau\|_{sp}$ the expression

$$\|\Lambda_\tau\|_{sp}=\hat{k}(0)+\frac{\hat{k}''(0)}{2}\frac{\pi^2}{\tau^2}-o(\tau^{-2}), \tag{5.5}$$

which agrees with [15].

II. *Hilbert's kernel.* Hardy, Littlewood and Pólya [7] considered the following bilinear form

$$\left|\sum_{m,n=1}^{\infty}\frac{a_m b_n}{m+n}\right|<\pi\left(\sum_{n=1}^{\infty}b_n^2\right)^{1/2}\left(\sum_{m=1}^{\infty}a_m^2\right)^{1/2}. \tag{5.6}$$

H. S. Wilf and N. G. de Bruijn [16] wished to find the best possible constant for which

$$\left|\sum_{m,n=1}^{N} \frac{a_m b_n}{m+n}\right| < M_N\left(\sum_{n=1}^{N} b_n^2\right)^{1/2}\left(\sum_{m=1}^{N} a_m^2\right)^{1/2}. \tag{5.7}$$

They showed [16, p. 32] that determining the constant M_N was equivalent to determining the spectral radius of an integral operator of the form

$$\Lambda_\tau f(x) = \int_{-\tau/2}^{\tau/2} k(x-y)f(y)\,dy \quad \text{where } k(x) = \frac{1}{2}\operatorname{sech}\frac{x}{2}. \tag{5.8}$$

It can be shown that

$$k_3(x) = 4\sum_{i=0}^{\infty} (-1)^i e^{-(2i+1)x/2}/(2i+1)^2, \tag{5.9}$$

and condition (3.1) leads to

$$\mu^* = \left(\max_{[0,1]} \frac{2\int_0^x \frac{\tan^{-1} s}{s}\,ds}{\tan^{-1} x}\right)^{-1}. \tag{5.10}$$

As in the previous example, p can be obtained as a bounded continuous function of τ.

Observing that $\hat{k}(\xi) = \pi \operatorname{sech} \pi\xi$, we can get the following asymptotics on $\|\Lambda_\tau\|_{sp}$,

$$\|\Lambda_\tau\|_{sp} = \pi - \pi^5/2\tau^2 + o(\tau^{-2}), \tag{5.11}$$

which agrees with [15] and [16].

III. *The Gaussian kernel.* $k(x) = \exp(-x^2)$. For this kernel, we have

$$k_3(x) = \int_x^\infty \int_t^\infty e^{-s^2}\,ds\,dt. \tag{5.12}$$

The condition that $(d/dx)(e^{\mu x}k_3(x)) \leqq 0$ leads to

$$\mu^* = \left(\sup_{x\geqq 0} \frac{\int_x^\infty \int_t^\infty e^{-s^2}\,ds\,dt}{\int_x^\infty e^{-s^2}\,ds}\right)^{-1}.$$

The ratio in parentheses is a bounded function of x, since it is a continuous function of x and goes to zero as $x \to \infty$. An easy calculation shows that

$$\hat{k}(\xi) = \sqrt{\pi}\, e^{-\xi^2/4},$$

and thus

$$\hat{k}\left(\frac{\pi}{\tau}\right) \leqq \|\Lambda_\tau\|_{sp} \leqq \hat{k}\left(\frac{\pi}{\tau + p(\tau)}\right) \tag{5.13}$$

leads to

$$\|\Lambda_\tau\|_{sp} = \sqrt{\pi} - \frac{\pi^{3/2}}{4\tau^2} + o(\tau^{-2}). \tag{5.14}$$

Acknowledgments. The author wishes to express his gratitude to Professor G. Milton Wing, Visiting Professor, Texas Tech University, for our many valuable discussions in the course of developing the results of this paper. He also wishes to express his appreciation to the referee for pointing out reference [16] and for his many valuable suggestions.

REFERENCES

[1] J. J. BOLMARCICH, *The behavior of the maximum value of finite sections of a class of bilinear forms*, J. Math. Anal. Appl., to appear.

[2] R. L. BOWDEN, W. GREENBERG AND P. F. ZWEIFEL, *Critical multigroup transport*, SIAM J. Appl. Math., 4 (1977), pp. 765–777.

[3] R. L. BOWDEN, S. SANCAKTAR AND P. F. ZWEIFEL, *Multigroup neutron transport*. I. J. Mathematical Phys., 17 (1976,) pp. 76–81.

[4] ———, *Multigroup neutron transport*. II. *Half range*, Ibid., 17 (1976), pp. 82–86.

[5] E. E. BURNISTON, T. W. MULLIKIN AND C. E. SIEWERT, *Steady state solutions in the two-group theory of neutron diffusion*, Ibid., 13 (1972), pp. 1461–1465.

[6] I. W. BUSBRIDGE, *The Mathematics of Radiative Transfer*, Cambridge University Press, London, 1960.

[7] G. H. HARDY, J. E. LITTLEWOOD AND G. PÓLYA, *Inequalities*, second ed., Cambridge University Press, London, 1964.

[8] S. KARLIN, *Positive operators*, J. Math. Mech., 8 (1959), pp. 907–937.

[9] M. G. KREIN, *Integral equations on a half-line with kernel depending upon the difference of the arguments*, Amer. Math. Soc. Transl., 22 (1962), pp. 163–288.

[10] J. T. KRIESE, C. E. SIEWERT AND Y. YENER, *Two-group critical problems for slabs and spheres in neutron transport theory*, Nuclear Sci. Engng., 50 (1973), pp. 3–9.

[11] A. LEONARD AND T. W. MULLIKIN, *Integral equations with difference kernels on finite intervals*, Trans. Amer. Math. Soc., 116 (1965), pp. 465–73.

[12] T. W. MULLIKIN, *Neutron branching processes*, J. Math. Anal. Appl., 3 (1961), pp. 507–25.

[13] T. W. MULLIKIN AND DEAN VICTORY, *N-group neutron transport theory: A criticality problem in slab geometry*, J. Math. Anal. Appl., 3 (1977), pp. 605–630.

[14] F. RELLICH, *Perturbation Theory of Eigenvalue Problems*, Gordon and Breach, New York, 1969.

[15] H. WIDOM, *Extreme eigenvalues of N-dimensional convolution operators*, Trans. Amer. Math. Soc., 106 (1963), pp. 391–414.

[16] H. S. WILF, *Finite Sections of Some Classical Inequalities*, Springer Verlag, New York 1970.

[17] G. MILTON WING, private communication.

SIAM J. MATH. ANAL.
Vol. 9, No. 3, June 1978

Copyright © 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0903-0012$1.00/0

GLOBAL VARIATION CRITERIA FOR THE L_2-STABILITY OF NONLINEAR TIME VARYING SYSTEMS*

Y. V. VENKATESH†

Abstract. In the framework of the positive operator theory of Zames, multiplier (general causal + anticausal) function form L_2-stability criteria are derived for a class of nonlinear time varying feedback systems represented by a time invariant stable part $\mathcal{G}$ in feedback with a nonlinear time varying gain $k(t)\varphi(\cdot)$. The criteria involve an upper global bound on the positive lobes of the normalized rate of variation, $\theta(t)=(dk/dt)/k$, and, simultaneously, a lower global bound on the negative lobes of $\theta(t)$; a trade-off, which is desirable in practice, between the two bounds is permitted. Finally, a method is described for partially freeing the stability conditions from multiplier dependence, but the problem of deriving explicit geometric stability conditions is still unsolved.

1. Introduction. Consider the feedback system illustrated in Fig. 1, where $\mathcal{G}$ is a time invariant linear operator, $\varphi(\cdot)$ is a memoryless monotone (or odd monotone) nonlinearity, and $k(t)$ is a time varying gain. We represent the system by the following equation:

$$\begin{aligned} v(t)&=x(t)-k(t)\varphi(y(t)),\\ y(t)&=(\mathcal{G}v)(t)=\sum_{i=1}^{\infty} g_i v(t-\tau_i)+\int_0^{\infty} g(t)v(t-\tau)\,d\tau \end{aligned} \tag{1}$$

for all $t \geqq 0$. Here $x(\cdot)$, $v(\cdot)$ and $y(\cdot)$ are respectively the input to the system, the error signal and output of the system. For assumptions on the components of (1), see § 2.

We derive L_2-stability[1] conditions in terms of the frequency response of $\mathcal{G}$ and a suitable multiplier function. The derivation is based on the theory of positive operators on a Hilbert space but in view of the integral equation formulation (1), the simpler "energy balance" argument [1] is employed.

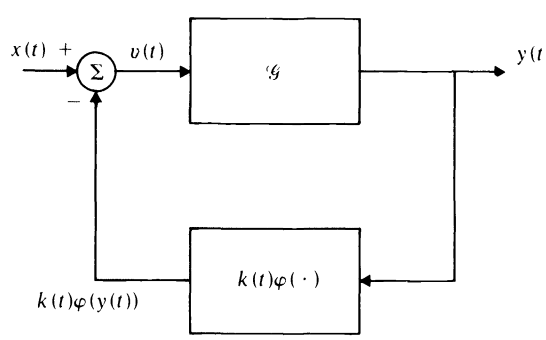

FIG. 1

The problem of stability of feedback systems with a single time varying nonlinearity was initially considered by Zames [2a], Sandberg [3] and others. A well-known

* Received by the editors July 24, 1975, and in final revised form November 22, 1976.

† Department of Electrical Engineering, Indian Institute of Science, Bangalore 560012, India. This work was supported by the Alexander von Humboldt Foundation, West Germany while the author was at Institut für Regelungs-und Steuerungssysteme, Universität Karlsruhe, West Germany.

[1] See Definition 4 in § 2.

criterion is the circle theorem which employs no more information than the range of values of the time varying nonlinearity. A more flexible result for systems with a monotone $\varphi(\cdot)$ was derived by Zames [2b] by employing an RC-multiplier and restricting the rate of variation of $k(t)$. For a reference to other contributions, see the brief survey in [4a]. More recently, for linear systems, Sundaresan and Thathachar [5a, b] established stability conditions involving upper and lower global bounds on $\theta(t)=(dk/dt)/k$ in the manner of Freedman and Zames [1]. See also Freedman [6]. But for nonlinear time varying systems, the following problem was left open.

Problem. Derive stability conditions in terms of lower and upper global bounds on $\theta(t)$ for the nonlinear time varying system represented by (1).

In fact, a footnote in [5a] conjectures the nonexistence of a solution to this problem.

The main result of the paper is Theorem 1 (§ 2) which gives L_2-stability conditions in terms of a multiplier function (containing causal and anticausal terms) and lower and upper global bounds on $\theta(t)$, a trade-off between the two bounds being permitted. The proof of the Theorem 1 is based on two lemmas (§ 3): the first lemma concerns the positivity of two operators one of which is linear time invariant and the other nonlinear time varying; the second lemma deals with the time varying gain factorization more general than that of Freedman and Zames [1]. Lemmas 1 and 2, believed to be of interest in their own right, are the main contributions of the paper. Once these lemmas are established, the positivity theorem of Zames [2a] along with the lemma on factorization of convolution operators defined on $(-\infty,\infty)$ [7] directly leads to the stability conditions. But in the interest of completeness and in view of the integral equation formation (1), the simpler energy balance argument is used (§ 3) to prove the main stability result. Finally, an attempt is made (§ 4) to free the stability conditions from dependence on the multiplier function.

2. Preliminaries, assumptions and statement of the main result.

DEFINITION 1. Let $L_2[0,\infty)$ be the linear space of real valued functions $x(\cdot)$ on $[0,\infty)$ with the property that

$$\int_0^\infty |x(t)|^2\,dt<\infty.$$

Let $L_2[0,\infty)$ be normed with the norm

$$\|x(\cdot)\|=\left(\int_0^\infty |x(t)|^2\,dt\right)^{1/2}.$$

The definition of the *extended space* L_{2e} is introduced via the notion of a truncated function $x_T(\cdot)$.

DEFINITION 2. For any real valued function $x(\cdot)$ on $[0,\infty)$ and any $T\geqq 0$, let $x_T(\cdot)$ denote the *truncated function* defined by

$$x_T(t)=\begin{cases}x(t) & \text{for } t\leqq T,\\ 0 & \text{for } t>T.\end{cases}$$

DEFINITION 3. Let L_{2e} be the space of those real valued functions $x(\cdot)$ on $[0,\infty)$ whose truncations $x_T(\cdot)$ belong to $L_2[0,\infty)$ for all $T\geqq 0$.

DEFINITION 4. The system described by (1) is L_2-stable if $v\in L_2\,[0,\infty)$ for $x\in L_2[0,\infty)$, and an inequality of the type $\|v\|\leqq\text{const.}\,\|x\|$ holds.

Concerning the feedback system described by (1), we make the following assumptions:

A1. $x(\cdot)$ is in $L_2[0, \infty)$.

A2. $v(\cdot)$ and $y(\cdot)$ are in L_{2e}.

A3. $g(\cdot)$ is a real valued element of $L_1[0, \infty)$, i.e., $\int_0^\infty |g(t)|\, dt < \infty$. There is a constant $\varepsilon_0 > 0$ such that $g(t) \exp(\varepsilon_0 t)$ is also in $L_1[0, \infty)$.

A4. $\{g_i\}$ is a sequence in l_1, i.e., $\sum_i |g_i| < \infty$. $\{\tau_i\}$ is a sequence in $[0, \infty)$.

A5. $k(\cdot)$ is assumed to be absolutely continuous on the interval $[0, \infty)$ and takes values in $[\varepsilon, \infty)$ for some constant $\varepsilon > 0$.

A6. $\varphi(\cdot)$ is a real valued function on $(-\infty, \infty)$. Further, $\varphi(0) = 0$, and $\varphi(\cdot)$ is a monotone nondecreasing, i.e., $(\sigma_1 - \sigma_2)(\varphi(\sigma_1) - \varphi(\sigma_2)) \geqq 0$ for all σ_1 and σ_2. There exist constants $q_1, q_2 > 0$ with $q_1 < q_2$ such that $q_1\sigma^2 \leqq \varphi(\sigma)\sigma \leqq q_2\sigma^2$ for all $\sigma \neq 0$. This class of nonlinearities is denoted by C. Let $G(j\omega)$ denote the frequency function of $\mathscr{G}$, i.e., $G(j\omega) = \sum_{i=1} g_i \exp(-j\omega\tau_i) + \int_0^\infty g(t) \exp(-j\omega t)\, dt$.

L₂-*Stability problem*. Find conditions on $k(t)$ and $G(j\omega)$ which ensure that $v(\cdot)$ is in $L_2[0, \infty)$ with $\|v\| \leqq \text{const.}\|x\|$.

A solution to the stability problem involves some additional definitions.

DEFINITION 5. Let P denote the class of operators $\mathscr{Z} \in L_{2e} \to L_{2e}$ satisfying an equation of the type

$$(2) \qquad (\mathscr{Z}x)(t) = x(t) + \sum_{i=1}^{\infty} z_i x(t - \sigma_i) + \sum_{i=1}^{\infty} z_i' x(t + \sigma_i') + \int_{-\infty}^{\infty} z(\tau)x(t - \tau)\, d\tau$$

where the sequences $\{z_i\}$ and $\{z_i'\}$ are in l_1, i.e., $\sum_{i=1}^{\infty} (|z_i| + |z_i'|) < \infty$; sequences $\{\sigma_i\}$, $\{\sigma_i'\}$ are in $[0, \infty)$; $z(\cdot)$ is a real valued function on $(-\infty, \infty)$, and is in $L_1(-\infty, \infty)$, i.e., $\int_{-\infty}^{\infty} |z(t)|\, dt < \infty$. The frequency function $Z(j\omega)$ is given by

$$Z(j\omega) = 1 + \sum_{i=1}^{\infty} z_i \exp(-j\omega\sigma_i) + \sum_{i=1}^{\infty} z_i' \exp(j\omega\sigma_i') + \int_{-\infty}^{\infty} z(\tau) \exp(-j\omega\tau)\, d\tau.$$

DEFINITION 6. Let

$$\theta(t) = \left(\frac{dk}{dt}\Big/ k\right),$$

$$\theta^+(t) = \begin{cases} \theta(t) & \text{for all } \theta(t) > 0, \\ 0 & \text{for all } \theta(t) \leqq 0; \end{cases}$$

and

$$\theta^-(t) = \begin{cases} \theta(t) & \text{for all } \theta(t) < 0, \\ 0 & \text{for all } \theta(t) \geqq 0. \end{cases}$$

Evidently, $\theta(t) = \theta^+(t) + \theta^-(t)$.

DEFINITION 7. Let

$$(3) \qquad \delta_s = \sup_{\substack{y \\ y \neq 0}} \left(\int_0^y \varphi(w)\, dw / \varphi(y)y \right),$$

$$(4) \qquad \delta_i = \inf_{\substack{y \\ y \neq 0}} \left(\int_0^y \varphi(w)\, dw / \varphi(y)y \right).$$

Note that for $\varphi(\cdot) \in C$, $0 < \delta_s \leqq 1$.

Throughout the rest of the paper, Re denotes "the real part of."

The main result of the paper is the following theorem:

THEOREM 1. *If there exists an operator* $\mathcal{Z}$ *in* P *with* $z(\cdot)\leqq 0$, $z_i\leqq 0$ *and* $z_i'\leqq 0$ *for all* $i=1,2,\cdots$, *such that*

(*a*) *for some positive constants* ξ, ζ

$$\sum_{i=1}^{\infty}|z_i|\exp(\xi\sigma_i)+\sum_{i=1}^{\infty}|z_i'|\exp(\zeta\sigma_i')+\int_0^{\infty}|z(\tau)|\exp(\xi\tau)\,d\tau + \int_{-\infty}^{0}|z(\tau)|\exp(-\zeta\tau)\,d\tau\leqq 1/(1+\delta_s-\delta_i); \tag{5}$$

(b) Re $Z(j\omega-\varepsilon)G(j\omega-\varepsilon)\geqq\delta>0$ *for all* ω *in* $(-\infty,\infty)$ *and some positive constant* ε (*which, in view of hypothesis* (a) *and assumption* A3, *is less than* ε_0); *and*

(c) *for some positive constants* N_1 *and* N_2, *and for all finite* $T>0$ *and all* $t_0\geqq 0$,

$$\frac{1}{T}\int_{t_0}^{t_0+T}\theta^+(t)\,dt\leqq N_1; \tag{6}$$

$$-N_2\leqq\frac{1}{T}\int_{t_0}^{t_0+T}\theta^-(t)\,dt \tag{7}$$

and in addition one of the following two sets of inequalities is satisfied:

Set 1.

$$\text{(i)}\quad \lim_{T\to\infty}\frac{1}{T}\int_{t_0}^{t_0+T}\theta^+(t)\,dt\leqq\xi \tag{8}$$

and

$$\text{(ii)}\quad \lim_{T\to\infty}\frac{1}{T}\int_{t_0}^{t_0+T}\theta^-(t)\,dt\geqq-\xi; \tag{9}$$

Set 2. (i) *for* $\xi<\zeta$,

$$\lim_{T\to\infty}\frac{1}{T}\int_{t_0}^{t_0+T}\theta^-(t)\,dt\geqq\xi-\zeta,\qquad \lim_{T\to\infty}\frac{1}{T}\int_{t_0}^{t_0+T}\theta^+(t)\,dt=0, \tag{10}$$

(ii) for $\xi>\zeta$,

$$\lim_{T\to\infty}\frac{1}{T}\int_{t_0}^{t_0+T}\theta^+(t)\,dt\leqq\xi-\zeta,\qquad \lim_{T\to\infty}\frac{1}{T}\int_{t_0}^{t_0+T}\theta^-(t)\,dt=0 \tag{11}$$

and

(iii) *for* $\zeta=\xi$, $\theta(t)$ *unrestricted*;

then the feedback system represented by (1) *is* L_2-*stable*.

Remark 1. Note that, apart from the flexibility in the choice of restriction on $\theta^+(t)$ and $\theta^-(t)$ by picking up Set 1 or Set 2 conditions, there is an additional freedom in the choice of ξ and ζ while satisfying the time domain inequality (5).

3. Principal lemmas and proof of the main result. Before we state the lemmas, we need the following definition.

DEFINITION 8. Let $\mathcal{K}$ be the class of absolutely continuous real valued functions $k(\cdot)$ on $[0, \infty)$ with each $k(\cdot)$ having constants $\underline{k}>0$ and $\bar{k} \geqq \underline{k}$ for which $\underline{k} \leqq k(t) \leqq \bar{k}$ for all $t \geqq 0$.

The following lemma concerns the positivity of two operators in cascade, one of them linear time invariant and the other nonlinear time varying. Zames and Falb [7, Lemma 8, Proposition 1] considered the positivity of a linear time invariant operator in cascade with a time invariant nonlinearity.

LEMMA 1. *If* (a) *the operator* $\mathcal{Z}$ *belongs to P with* $z(\cdot) \leqq 0$, $z_i \leqq 0$ *and* $z_i' \leqq 0$ *for all* $i = 1, 2, \cdots$; (b) *for some nonnegative constants* ξ *and* ζ, *inequality* (5) *is satisfied*; *and* (c) *for these values of* ξ *and* ζ, *with* $f(\cdot) \in \mathcal{K}$, $f(t)k(t) \exp(-\xi t)$ *is nonincreasing and* $f(t)k(t) \exp(\zeta t)$ *nondecreasing for all* $t \geqq 0$, *then the following inequality holds*:

$$\int_0^T f(t)(\mathcal{Z}x)(t)k(t)\varphi(x(t))\,dt \geqq 0 \tag{12}$$

for all x in the domain of $\mathcal{Z}$ *and for all* $T \geqq 0$.

Proof. See Appendix A.

Remark 2. For odd monotone $\varphi(\cdot)$, Lemma 1 is applicable without, in hypothesis (a) of Lemma 1, the nonpositivity constraint on $z(\cdot)$, z_i and z_i' for all $i = 1, 2, \cdots$.

The next lemma deals with the factorization of time varying gains.

LEMMA 2. *If there exists a time multiplier function* $f(\cdot)$ *in* $\mathcal{K}$ *satisfying hypothesis* (c) *of Lemma* 1 *for some positive constants* ξ *and* ζ, *then hypothesis* (c) *of Theorem* 1 *holds*.

Proof. See Appendix B.

Remark 3. Lemma 2 seems to be a significant generalization of the lemma of Freedman and Zames [1, Lemma 4].

Based on Lemmas 1 and 2, Theorem 1 may now be proved.

Proof of Theorem 1. Consider the integral, for any $T>0$,

$$\rho(T) = \int_0^T f(t)x(t)(\mathcal{Z}\mathcal{G}v)(t)\,dt \tag{13}$$

which when we use (1) becomes

$$\rho(T) = \int_0^T f(t)v(t)(\mathcal{Z}\mathcal{G}v)(t)\,dt + \int_0^T f(t)k(t)\varphi(y(t))(\mathcal{Z}y)(t)\,dt. \tag{14}$$

The first integral on the right hand side of (14) is, for some $\varepsilon>0$,

$$\int_0^T f(t)v(t)(\mathcal{Z}\mathcal{G}v)(t)\,dt = \int_0^T f(t)\exp(-2\varepsilon t)v(t)(\mathcal{Z}\mathcal{G}v)(t)\exp(2\varepsilon t)\,dt.$$

Suppose we choose $f \in \mathcal{K}$ so that $f(t)\exp(-2\varepsilon t)$ is nonincreasing and $\varepsilon < \varepsilon_0$ (see hypothesis (b) of Theorem 1). Then we can invoke the second mean value theorem (Hobson [8, p. 618]). According to this, there is a point T' in $[0, T]$ for which

$$\int_0^T f(t)v(t)(\mathcal{Z}\mathcal{G}v)(t)\,dt = f(0)\int_0^{T'} v(t)(\mathcal{Z}\mathcal{G}v)(t)\exp(2\varepsilon t)\,dt$$

where $f(0)>0$ in view of our choice $f \in \mathcal{K}$. By hypothesis (b) of Theorem 1, $\operatorname{Re} Z(j\omega-\varepsilon)G(j\omega-\varepsilon) \geqq \delta > 0$ for all ω in $(-\infty, \infty)$. Hence by Parseval's theorem

$$\int_0^T f(t)v(t)(\mathcal{Z}\mathcal{G}v)(t)\,dt \geqq \delta_1\|v_T\|^2 \tag{15}$$

for some constant $\delta_1 > 0$.

Now consider the second integral on the right hand side of (14). By virtue of Lemmas 1 and 2, this integral is nonnegative if hypotheses (a) and (c) of the theorem statement are satisfied. Consequently, from (15), (14) and from an application of Parseval's theorem to the integral of (13), we have

$$\delta_1 \|v_T\|^2 \leqq \int_0^T f(t)x(t)(\mathscr{Z}\mathscr{G}v)(t)\,dt \tag{16}$$
$$\leqq \bar{f}\|x_T\|\,\|v_T\| \sup_{-\infty<\omega<\infty} |Z(j\omega)G(j\omega)|;$$

but $\sup_{-\infty<\omega<\infty} |Z(j\omega)G(j\omega)|$ is finite by virtue of the assumptions on $\mathscr{Z}$ and $\mathscr{G}$. Hence with constant $A = \bar{f} \sup_{-\infty<\omega<\infty} |Z(j\omega)G(j\omega)|$, we get from (16) the inequality

$$\delta_1 \|v_T\| \leqq A \|x_T\|$$

which is valid for all $T > 0$. The theorem is proved.

Remark 4. Theorem 1 applies to nonlinear systems with odd monotone nonlinearity under less restrictive assumptions on the multiplier: the nonpositivity constraint on $z(\cdot)$ in Theorem 1 is no longer necessary. (See Remark 2 above.) Note that, when $\xi = \zeta$, in hypothesis (c), choice of Set 2 imposes no constraint on $\theta(t)$. In this form Theorem 1 is believed to be a generalization of the circle criterion [2] for $\varphi(\cdot)$ monotone.

4. Construction of a multiplier function $Z(j\omega)$. It is assumed that $G(j\omega-\varepsilon)$ for some constant $\varepsilon > 0$ satisfies the Nyquist criterion: that is, $\arg G(j\omega-\varepsilon)$ lies in the interval $(-\pi, \pi)$ for $\omega \in (-\infty, \infty)$. Note that $\varepsilon < \varepsilon_0$ by assumption A3, § 2. If a multiplier function $Z(j\omega)$ is so chosen that $\arg Z(j\omega)$ and $\arg Z(j\omega)G(j\omega)$ lie in the interval $((-\pi/2)+\alpha,\ (\pi/2)-\alpha)$ for all $\omega \in (-\infty, \infty)$ and for some $\alpha > 0$, then $\arg Z(j\omega-\varepsilon)G(j\omega-\varepsilon)$ also lies in $((-\pi/2)+\alpha,\ (\pi/2)-\alpha)$ for some positive small constants ε and α. Thus it remains to verify hypothesis (a) and, with the maximum value of constants ξ and ζ so obtained, to satisfy hypothesis (c) of the theorem statement.

Define

$$\psi(\omega) = \begin{cases} (\pi/2) - \arg G(j\omega) & \text{for } \arg G(j\omega) \geqq 0, \\ (-\pi/2) - \arg G(j\omega) & \text{for } \arg G(j\omega) < 0. \end{cases}$$

Then $\arg Z(j\omega)$ lies in the band formed by two functions ψ_1 and ψ_2 defined as follows:

$$\begin{aligned} \psi_1(\omega) &= ((-\pi/2)+\alpha,\ \psi(\omega)) - \alpha \quad && \text{for } \arg G(j\omega) \geqq 0, \\ \psi_2(\omega) &= (\psi(\omega)+\alpha,\ (\pi/2)-\alpha) && \text{for } \arg G(j\omega) < 0 \end{aligned} \tag{17}$$

for some $\alpha > 0$.

Remark 5. The multiplier used by Freedman and Zames [1] in their final stability results is a causal function, and the construction is based on the assumption that $\arg G(j\omega)$ approaches zero as $|\omega| \to \infty$. This assumption excludes consideration of systems with arguments tending to $\pm\pi/2$ or $\pm\pi$ as $|\omega| \to \infty$.

We now examine two extreme cases.

Case 1. It is assumed that there is a minimum finite frequency W beyond which $|\arg G(j\omega)| < \pi/2$ for all $|\omega| > W$.

Case 2. It is assumed that $\arg G(j\omega)$ tends to $-\pi$ as $\omega \to \infty$.

Construction of multiplier function: Case 1. Let $\Phi_0(\omega)$ be a real valued continuous almost everywhere differentiable function chosen in the band (17) for $|\omega| < W$, and outside the interval $(-W, W)$, let $\Phi_0(\omega) = 0$. If we set $\Phi_0(\omega) = -\Phi_0(\omega)$, it becomes an odd function of ω. It can be concluded that $\Phi_0(\omega)$ and $d\Phi_0/d\omega$ are in $L_2(-\infty, \infty)$ and hence the inverse (limit-in-the-mean) Fourier transform $\phi_0(t)$ of $j\Phi_0(\omega)$ is real, odd and in $L_1(-\infty, \infty)$. Define $\phi_e(t)$ on $(-\infty, \infty)$ as follows:

$$\phi_e(t) = \begin{cases} \gamma\phi_0(t) & \text{for } t \geqq 0, \\ -\gamma\phi_0(t) & \text{for } t < 0 \end{cases} \tag{18}$$

for some constant γ. Then $\phi_e(t)$ is real, even and is in $L_1(-\infty, \infty)$. Its Fourier transform $\Phi_e(\omega)$ is even. Now let

$$\begin{aligned} \lambda(t) = \phi_e(t) + \phi_0(t) &= (\gamma+1)\phi_0(t) \quad \text{for } t \geqq 0, \\ &= (-\gamma+1)\phi_0(t) \quad \text{for } t < 0. \end{aligned} \tag{19}$$

Its Fourier transform is given by

$$\Lambda(j\omega) = \Phi_e(\omega) + j\Phi_0(\omega). \tag{20}$$

Further, let $(\lambda * \lambda)(t)$ denote the convolution product:

$$(\lambda * \lambda)(t) = \int_{-\infty}^{\infty} \lambda(\tau)\lambda(t-\tau)\, d\tau$$

and let the subscript 1 in a norm denote the L_1-norm as for instance in

$$\|\lambda\|_1 = \int_{-\infty}^{\infty} |\lambda(t)|\, dt.$$

We now generate the multiplier function as follows:

$$\begin{aligned} Z(j\omega) &= 1 + \Lambda(j\omega) + (\Lambda^2(j\omega)/2!) + \cdots + (\Lambda^n(j\omega)/n!) + \cdots \\ &= \exp(\Lambda(j\omega)). \end{aligned} \tag{21}$$

For details on convergence see [1, Lemma 2]. It is easy to verify that $\arg G(j\omega) = \Phi_0(\omega)$ and hence in view of the construction of $\Phi_0(\omega)$ so as to lie in the band (17), hypothesis (b) of Theorem 1 is satisfied. It remains to verify hypothesis (c) of the theorem. To this end, let

$$\begin{aligned} z_1(t) &= \text{inverse Fourier transform of } \exp(\Lambda(j\omega)) - 1 \text{ for } t \geqq 0, \\ z_2(t) &= \text{inverse Fourier transform of } \exp(\Lambda(j\omega)) - 1 \text{ for } t < 0. \end{aligned} \tag{22}$$

Find ξ and $\zeta > 0$ such that

$$\|z_1 \exp(\xi t)\|_1 + \|z_2 \exp(-\zeta t)\|_1 \leqq 1/(1+\delta_s-\delta_i) \tag{23}$$

after suitably choosing γ in (18).

Remark 6. For $\gamma = 1$ the multiplier $Z(j\omega)$ is the one used by Freedman and Zames [1] in which some additional constraints are imposed. The multiplier of [1] is a causal function, i.e., $z_2(t) = 0$ for $t < 0$. The time domain inequality (23) becomes

$$\|\phi_0 \exp(\xi t)\|_1 \leqq \tfrac{1}{2} \log(1 + (1/(1+\delta_s-\delta_i)))$$

with the left hand side norm taken in the interval $[0, \infty)$. Further, note that $\zeta = \infty$, implying thereby that the behavior of the negative lobes of $\theta(t)$ is unrestricted. In a similar manner, for $\gamma = -1$, $z_1(t) = 0$ for $t \geqq 0$ and $\xi = \infty$. This implies that the behavior of the positive lobes of $\theta(t)$ is unrestricted. Consequently, for those extreme cases, in hypothesis (c) of the theorem, only the Set 1 bounds on $\theta(t)$ are meaningful. For values of γ in $(-1, 1)$, from inequality (23) there is obviously a trade-off between ξ and ζ: the larger the value of ξ, the smaller is the value of ζ and vice versa. Note also that no practical advantage is gained by choosing $|\gamma|>1$.

The following theorem is an alternative version of Theorem 1 for Case 1.

THEOREM 2. *If* $|\arg G(j\omega)|<\pi/2$ *for all* $|\omega|>W$, *a finite number, and there exists a function* $\Phi_0(\omega)$ *in the band* (17) *within the interval* $(-W, W)$ *and equal to zero outside this interval, such that, with* $\phi_0(t)=$ *inverse Fourier transform of* $j\Phi_0(\omega)$, (a) $z_1(t)$ *and* $z_2(t)$, *defined by* (22) *and* (21), *are nonpositive*; (b) *inequality* (23) *holds for some* $\xi, \zeta>0$; *and* (c) *hypothesis* (c) *of Theorem* 1 *is satisfied*; *then the feedback system represented by* (1) *is* L_2*-stable.*

Remark 7. We note that Theorem 2 holds for the feedback system with an odd monotonic nonlinearity $\varphi(\cdot)$ by removing the nonpositivity restriction on $z_1(t)$ and $z_2(t)$, and thus a simpler geometric interpretation can be given:

COROLLARY. *With* $\Phi_0(\omega)$ *chosen in the band* (17), $(\tan \Phi_0(\omega))$ *is finite in* $(-W, W)$ *and zero outside this interval. Further* $(\tan \Phi_0(\omega))$ *and its derivatives are in* $L_2(-\infty, \infty)$ *and hence the inverse (limit-in-the-mean) Fourier transform,* $z_e(t)$, *of* $(j \tan \Phi_0(\omega))$ *is real, odd and is in* $L_1(-\infty, \infty)$. *If there exists a positive* ξ *such that*

$$\int_0^{\infty} |z_e(t)| \exp(\xi t)\, dt \leqq 1/(2(1+\delta_s-\delta_i))$$

and for the value of ξ *so obtained, hypothesis* (c) *of Theorem* 1 *is satisfied for* $\xi=\zeta$, *then the feedback system represented by* (1) *with an odd monotonic* $\varphi(\cdot)$ *is* L_2*-stable.*

However, in general, an interpretation of inequality (23) in terms of $\Phi_0(\omega)$ and its derivative, as attempted by Freedman [6b] for the case of time invariant systems and causal multipliers, seems to be difficult and constitutes an open problem.

Now we consider the case of $\arg G(\omega)$ tending to $-\pi$ as $|\omega|\to\infty$ (Case 2) for constructing a multiplier function.

Construction of a multiplier function: Case 2. Find the minimum value of the constant γ such that $\arg((1+j\gamma\omega)G(j\omega))$ lies in the band $[0, \pi/2)$ for $|\omega|>W$, a finite number. Let $Z(j\omega)=(\exp \Lambda(j\omega)+j\gamma\omega)$ with $\Lambda(j\omega)$constructed as for Case 1 but with $G(j\omega)$ replaced by $(1+j\gamma\omega)G(j\omega)$ for $|\omega|<W$. Because of the additional term in the multiplier, Theorem 1 and consequently Theorem 2 need a minor change which is given in Appendix C.

Conclusions. The Zames positive operator principle is used to derive interchangeable upper and global constraints on the rate of variation of the time varying gain for the L_2-stability of nonlinear time varying feedback systems which are not necessarily described by a differential equation. These results are more general than those available in the literature and constitute a solution to the problems left open in [5]. A geometric interpretation of the stability results of the paper is not complete as presented here and hence seems to deserve further investigation.

In view of the relative slackening of the pace of publications in the area of time varying system stability, one is tempted to conclude, following John von Neumann, that we have reached the baroque stage. But on closer investigation one soon discovers that (a) all the results available on stability when applied to Mathieu's or Hill's equation are incapable of reproducing the well-known classical stability boundaries, and (b) in spite of the attempts made by researchers in the field of instability, the instability counterpart of, say, the stability criterion of Freedman and Zames [1] for linear time varying systems is not known. See in this context the comments in Skoog [9, p. 93]. It should be a challenging problem to derive the instability counterpart of Theorem 1 of this paper.

Appendix A.

Proof of Lemma 1. We have

$$\int_0^T f(t)(\mathscr{L}x)(t)k(t)\varphi(x(t))\,dt$$

$$= \int_0^T f(t)k(t)\Big\{x(t)+\sum_{i=1}^{\infty} z_i x(t-\sigma_i)$$

$$+\int_0^{\infty} z(\tau)x(t-\tau)\,d\tau+\sum_{i=1}^{\infty} z_i' x(t+\sigma_i')+\int_{-\infty}^{0} z(\tau)x(t-\tau)\,d\tau\Big\}\varphi(x(t))\,dt$$

$$\text{(A.1)}\qquad = \int_0^T f(t)k(t)\,e^{-\xi t}\Big\{\tfrac{1}{2}x(t)\,e^{\xi t}+\sum_{i=1}^{\infty} z_i\,e^{\xi t}x(t-\sigma_i)$$

$$+e^{\xi t}\int_0^{\infty} z(\tau)x(t-\tau)\,d\tau\Big\}\varphi(x(t))\,dt$$

$$+\int_0^T f(t)k(t)\,e^{\zeta t}\Big\{\tfrac{1}{2}x(t)\,e^{-\zeta t}$$

$$+\sum_{i=1}^{\infty} z_i'\,e^{-\zeta t}x(t+\sigma_i')+e^{-\zeta t}\int_{-\infty}^{0} z(\tau)x(t-\tau)\,d\tau\Big\}\varphi(x(t))\,dt$$

where ξ, ζ are nonnegative constants.

Since $f(t)k(t)\,e^{-\xi t}$ is nonincreasing, by the second mean value theorem, there is a point T' in $[0, T]$ for which the first integral on the right hand side of (A.1) becomes

$$f(0)k(0)\Big\{\int_0^{T'}\Big[\tfrac{1}{2}x(t)\,e^{\xi t}+\sum_{i=1}^{\infty}(z_i\,e^{\xi\sigma_i})(x(t-\sigma_i)\,e^{\xi(t-\sigma_i)})$$

$$+\int_0^{\infty}(z(\tau)\,e^{\xi\tau})x(t-\tau)\,e^{\xi(t-\tau)}\,d\tau\Big]\varphi(x(t))\,dt\Big\}$$

which, on interchange of the integration and summation operators (assuming the validity of such an interchange) assumes the form

$$\text{(A.2)}\qquad f(0)k(0)\Big\{\int_0^{T'}\tfrac{1}{2}x(t)\,e^{\xi t}\varphi(x(t))\,dt+\sum_{i=1}^{\infty} z_i\,e^{\xi\sigma_i}\int_0^{T'} x(t-\tau_i)\,e^{\xi(t-\sigma_i)}\varphi(x(t))\,dt$$

$$+\int_0^{\infty} z(\tau)\,e^{\xi\tau}\,d\tau\int_0^{T'} e^{\xi(t-\tau)}x(t-\tau)\varphi(x(t))\,dt\Big\}.$$

By virtue of Lemma 2 of [4b], it can be shown that the expression (A.2) is nonnegative if

$$\sum_{i=1}^{\infty}|z_i|\,e^{\xi\sigma_i}+\int_0^{\infty}|z(\tau)|\,e^{\xi\tau}\,d\tau \leqq \frac{1}{2(1+\delta_s-\delta_i)}.$$

As regards the second integral on the right hand side of (A.1), we note that $f(t)k(t)\,e^{\xi t}$ is given as nondecreasing. Hence by the second mean value theorem, there is a point T'' in $[0, T]$ such that the second integral on the right hand side of (A.1) becomes:

$$f(T)k(T)\,e^{\zeta T}\left\{\int_{T''}^{T}\left[\tfrac{1}{2}x(t)\,e^{-\zeta t}+\sum_{i=1}^{\infty} z_i'\,e^{\zeta\sigma_i'}x(t+\sigma_i)\,e^{-\zeta(t+\sigma_i)}\right.\right.$$
$$\left.\left.+\int_{-\infty}^{0} z(\tau)\,e^{-\zeta\tau}x(t-\tau)\,e^{-\zeta(t-\tau)}\,d\tau\right]\varphi(x(t))\,dt\right\}.$$

Proceeding in the manner given for the first integral on the right hand side of (A.1) and using Lemma 2 of [4b], we conclude that the second integral on the right hand side of (A.1) is nonnegative if

$$\sum_{i=1}^{\infty}|z_i'|\,e^{\zeta\sigma_i'}+\int_{-\infty}^{0}|z(\tau)|\,e^{-\zeta\tau}\,d\tau \leqq \frac{1}{2(1+\delta_s-\delta_i)}.$$

The lemma is proved.

Appendix B.

Proof of Lemma 2. We have

$$-\zeta \leqq \theta(t)+\left(\frac{df}{dt}\Big/ f\right)\leqq \xi, \qquad t\in[0,\infty). \tag{B.1}$$

Let

$$\left(\frac{df}{dt}\Big/ f\right)=h_1(t)+h_2(t). \tag{B.2}$$

Then inequality (B.1) is satisfied by choosing

$$h_1(t)=\xi-\theta^+(t); \qquad h_2(t)=-\zeta-\theta^-(t).$$

Hence, from (B.2), for all $t_0 \geqq 0$ and all $t \geqq t_0$,

$$f(t)=f(t_0)\exp\left(\int_{t_0}^{t}(-\zeta-\theta^-(\tau)-\theta^+(\tau)+\xi)\,d\tau\right). \tag{B.3}$$

But $f(\cdot)$ is in $\mathcal{K}$ and hence, for some positive constants β_1 and β_2 with $\beta_1<\beta_2$,

$$\beta_1 \leqq \exp\left(\int_{t_0}^{t}(-\zeta-\theta^-(\tau)-\theta^+(\tau)+\xi)\,d\tau\right)\leqq \beta_2 \tag{B.4}$$

for all $t_0 \geqq 0$ and all $t \geqq t_0$.

The inequality (B.4) can be satisfied in any one of the following ways: For all $t_0 \geqq 0$ and all $t \geqq t_0$, and for some positive constants M_1 and M_2, and M_3 and M_4,

Case 1.

$$-\infty<-M_1 \leqq \int_{t_0}^{t}(-\zeta+\xi-\theta(\tau))\,d\tau \leqq M_2<\infty. \tag{B.5}$$

Case 2.

$$\text{(i)}\quad -\infty<-M_1 \leqq \int_{t_0}^{t}(\xi-\theta^-(\tau))\,d\tau \leqq M_2<\infty; \tag{B.6}$$

and

$$\text{(ii)}\quad -\infty<-M_3 \leqq \int_{t_0}^{t}(-\zeta-\theta^+(\tau))\,d\tau \leqq M_4<\infty. \tag{B.7}$$

Case 3.

(B.8) $$\text{(i)}\quad -\infty < -M_1 \leqq \int_{t_0}^{t} (-\zeta - \theta^-(\tau))\, d\tau \leqq M_2 < \infty;$$

and

(B.9) $$\text{(ii)}\quad -\infty < -M_1 \leqq \int_{t_0}^{t} (\xi - \theta^+(\tau))\, d\tau \leqq M_4 < \infty.$$

Case 4.

(B.10) $$\text{(i)}\quad -\infty < -M_1 \leqq \int_{t_0}^{t} (\xi - \zeta - \theta^-(\tau))\, d\tau \leqq M_2 < \infty;$$

and

(B.11) $$\text{(ii)}\quad -\infty < -M_3 \leqq \int_{t_0}^{t} (-\theta^+(\tau))\, d\tau \leqq M_4 < \infty.$$

Case 5.

(B.12) $$\text{(i)}\quad -\infty < -M_1 \leqq \int_{t_0}^{t} (-\theta^-(\tau))\, d\tau \leqq M_2 \leqq \infty;$$

and

(B.13) $$\text{(ii)}\quad -\infty < -M_3 \leqq \int_{t_0}^{t} (\xi - \zeta - \theta^+(\tau))\, d\tau \leqq M_4 < \infty.$$

Case 1. Inequality (B.5) can be reduced to

(B.14) $$\exp(-M_2 + (\xi - \zeta)(t - t_0)) \leqq \frac{k(t)}{k(t_0)} \leqq \exp(M_1 + (\xi - \zeta)(t - t_0))$$

for all $t_0 \geqq 0$ and all $t \geqq t_0$.

When $\xi > \zeta$ and when $\xi < \zeta$, we conclude from (B.14) that $k(\cdot) \notin \mathscr{K}$. Hence Case 1 is ruled out for $\xi \neq \zeta$. But when $\xi = \zeta$, inequality (B.14) merely implies that $k(\cdot) \in \mathscr{K}$ and hence no restriction on $\theta(t)$ is imposed.

Case 2. From inequality (B.6),

(B.15) $$-M_2 + \xi(t - t_0) \leqq \int_{t_0}^{t} \theta^-(\tau)\, d\tau \leqq M_1 + \xi(t - t_0)$$

for all $t_0 \geqq 0$ and all $t \geqq t_0$. But (B.15) is untenable because $\theta^-(\tau)$ is always negative and hence the left hand side of the inequality is violated for sufficiently large $t > t_0$. (The right hand side of (B.15) is, however, trivially satisfied). In an identical manner, inequality (B.7) is untenable. Hence Case 2 is ruled out.

Case 3. From inequalities (B.8) and (B.9) respectively,

(B.16) $$-M_2 - \zeta(t - t_0) \leqq \int_{t_0}^{t} \theta^-(\tau)\, d\tau \leqq M_1 - \zeta(t - t_0),$$

(B.17) $$-M_4 + \xi(t - t_0) \leqq \int_{t_0}^{t} \theta^+(\tau)\, d\tau \leqq M_3 + \xi(t - t_0)$$

for all $t_0 \geqq 0$ and all $t \geqq t_0$.

Hence by requiring that $\int_{t_0}^t \theta^+(\tau)\,d\tau$ and $\int_{t_0}^t |\theta^-(\tau)|\,d\tau$ be bounded for all finite $t \geqq t_0$, and

$$\lim_{T\to\infty} \frac{1}{T}\int_{t_0}^{t_0+T} \theta^+(\tau)\,d\tau \leqq \xi,$$

$$\lim_{T\to\infty} \frac{1}{T}\int_{t_0}^{t_0+T} \theta^-(\tau)\,d\tau \geqq -\zeta,$$

inequalities (B.16) and (B.17) are satisfied.

Case 4. From inequalities (B.10) and (B.11) respectively,

$$\text{(B.18)} \qquad -M_2+(\xi-\zeta)(t-t_0) \leqq \int_{t_0}^t \theta^-(\tau)\,d\tau \leqq M_1+(\xi-\zeta)(t-t_0),$$

$$\text{(B.19)} \qquad -M_4 \leqq \int_{t_0}^t \theta^+(\tau)\,d\tau \leqq M_3$$

for all $t_0 \geqq 0$ and all $t \geqq t_0$.

When $\xi > \zeta$, the left hand side of inequality (B.18) is violated (in view of the negative nature of $\theta^-(\tau)$) for sufficiently large $t > t_0$. Hence inequality (B.18) is incompatible for $\xi > \zeta$. Consequently, we require that $\xi \leqq \zeta$.

When $\xi = \zeta$, (B.18) gives

$$\text{(B.20)} \qquad -M_2 \leqq \int_{t_0}^t \theta^-(\tau)\,d\tau \leqq M_1$$

for all $t_0 \geqq 0$ and all $t \geqq t_0$. Inequalities (B.19) and (B.20) are definitely more restrictive than the waiving of all restrictions on $\theta(t)$ in Case 1 for $\xi = \zeta$. Hence we need consider only $\xi < \zeta$. As in Case 3, if $\int_{t_0}^t |\theta^-(\tau)|\,d\tau$ is bounded for all finite $t \geqq t_0$, and

$$\lim_{T\to\infty} \frac{1}{T}\int_{t_0}^{t_0+T} \theta^-(\tau)\,d\tau \geqq \xi-\zeta$$

then (B.18) holds.

Case 5. As in Case 4, inequalities (B.12) and (B.13) are respectively equivalent to requiring that

$$-M_2 \leqq \int_{t_0}^t \theta^-(\tau)\,d\tau \leqq M_1$$

for all $t_0 \geqq 0$ and all $t \geqq t_0$; and, for $\xi > \zeta$, $\int_{t_0}^t \theta^+(\tau)\,d\tau$ is bounded for all finite $t \geqq t_0$,

$$\lim_{T\to\infty} \int_{t_0}^{t_0+T} \theta^+(\tau)\,d\tau \leqq \xi-\zeta.$$

When $\xi < \zeta$, $k(\cdot) \notin \mathcal{K}$, and when $\xi = \zeta$, the constraint so obtained on $\theta(t)$ is stronger than necessary as explained in Case 4.

The lemma is proved.

Appendix C. Consider the integral

$$\rho_1(T) = \int_0^T f(t)k(t)\Big(x(t)+\gamma\frac{dx}{dt}\Big)\varphi(x(t))\,dt$$

where γ is a positive constant. This integral can be rewritten as

$$\rho_1(T)=\int_0^T f(t)k(t)\,e^{-\nu t}(x(t)\,e^{\nu t}+\nu\,e^{\nu t}(dx/dt))\varphi(x(t))\,dt \tag{C.1}$$

where ν is another positive constant which is to be suitably chosen.

If $f(t)k(t)\exp(-\nu t)$ is nonincreasing, there is a point T' in $[0, T]$ for which (C.1) takes the form

$$\rho_1(T)=f(0)k(0)\int_0^{T'}\left(x(t)\,e^{\nu t}+\gamma\,e^{\nu t}\frac{dx}{dt}\right)\varphi(x(t))\,dt. \tag{C.2}$$

On integrating by parts the portion of (C.2) containing the derivative of x, and simplifying, we get

$$\rho_1(T)=f(0)k(0)\int_0^{T'} e^{\nu t}\left\{x(t)\varphi(x(t))-\gamma\nu\int_0^{x(t)}\varphi(u)\,du\right\}dt$$
$$+f(0)k(0)\left\{e^{\nu t}\int_0^{x(t)}\varphi(u)\,du\right\}\Bigg|_0^{T'}$$

which on using (3) leads to the inequality

$$\rho_1(T)\geqq f(0)k(0)\int_0^{T'}(1-\gamma\nu\delta_s)x(t)\varphi(x(t))\,dt+f(0)k(0)\left\{e^{\nu t}\int_0^{x(t)}\varphi(u)\,du\right\}\Bigg|_0^{T'}.$$

Thus Lemma 2 remains valid with $\mathscr{X}$ replaced by $(\mathscr{X}+\gamma(d/dt))$ and the right hand side of inequality (5) by $(1-\gamma\nu\delta_s)/(1+\delta_s-\delta_i)$. But then, at the same time, ν should be so chosen that $f(t)k(t)\exp(-\nu t)$ and $f(t)k(t)\exp(-\xi t)$ are nonincreasing for values of ν and ξ which are almost equal to each other. Even though the choice of ν equal to ξ would be appropriate, it is observed that the right hand side of the inequality (5) is reduced correspondingly. The problem is then one of finding out an optimal ν satisfying inequality (5) with the modified right hand side and such that it is on a par with ξ.

Note further that

$$(1+j\gamma\omega+Z(j\omega))G(j\omega)=(1+j\gamma\omega)G(j\omega)\left(1+\frac{Z(j\omega)}{(1+j\gamma\omega)}\right)$$

and hence, in this case,

$$(\exp\Lambda(j\omega))-1=Z(j\omega)/(1+j\gamma\omega).$$

The changes to be made in the statement of Theorem 2 are: (i) Substitute the inverse transform of $(1+j\gamma\omega)((\exp\Lambda(j\omega))-1)$ for the inverse transform of $(\exp\Lambda(j\omega)-1)$; and (ii) replace the right hand side of (23) by $(1-\gamma\nu\delta_s)/(1+\delta_s-\delta_i)$.

Acknowledgment. The author is grateful to the Alexander von Humboldt Foundation, West Germany for the financial support and to Professor O. Föllinger for encouragement and interest. Thanks are extended to the expert referees for their valuable suggestions.

REFERENCES

[1] M. FREEDMAN AND G. ZAMES, *Logarithmic variation criteria for the stability of systems with time varying gains*, SIAM J. Control, 6 (1968), pp. 487–507.

[2a] G. ZAMES, *On the stability of nonlinear time varying feedback systems*, Proc. National Electronic Conf. (Chicago), vol. 20, 1964, pp. 725–730.

[2b] ———, *Nonlinear time varying feedback systems—conditions for L_∞-boundedness using conic operators on exponentially weighted spaces*, Proc. Third Allerton Conference, University of Illinois, Monticello, Il., 1965, pp. 460–471.

[3] I. W. SANDBERG, *A frequency domain condition for the stability of feedback systems containing a single time varying nonlinear element*, Bell System Tech. J., 43 (1964), pp. 1601–1608.

[4a] Y. V. VENKATESH, Noncausal multipliers for nonlinear system stability, IEEE Trans. Automatic Control, AC-15 (1970), pp. 195–204.

[4b] ———, *On the positivity of nonlinear time varying operators*, Ibid., AC-18 (1973), pp. 321–322.

[5a] M. K. SUNDARESAN AND M. A. L. THATHACHAR, *Time varying systems stability—interchangeability of the bounds on the logarithmic variation of gain*, Ibid., AC-18 (1973), pp. 405–407.

[5b] ———, *Average variation L_2-stability criteria for time-varying feedback systems—a unified approach*, Ibid., AC-19 (1974), pp. 427–429.

[6a] M. FREEDMAN, *L_2-stability of time varying systems*, SIAM J. Control, 6 (1968), pp. 559–578.

[6b] ———, *Phase function norm estimates for stability of systems with monotone nonlinearities*, Ibid., 10 (1972), pp. 99–111.

[7] G. ZAMES AND P. L. FALB, *Stability conditions for systems with monotone and slope restricted nonlinearities*, Ibid., 6 (1968), pp. 89–108.

[8] E. W. HOBSON, *The Theory of Functions of a Real Variable and an Introduction to the Theory of Fourier Series*, vol. 1, Cambridge University Press, London, 1927.

[9] R. A. SKOOG, *Positivity conditions and instability criteria for feedback systems*, SIAM J. Control, 12 (1974), pp. 83–98.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0001 $01.00/0

NONHOMOGENEOUS SECOND ORDER DIFFERENTIAL SYSTEMS WITH INTEGRAL BOUNDARY CONDITIONS*

S. C. TEFTELLER†

Abstract. This paper is concerned with second order, nonhomogeneous differential systems involving a parameter with integral boundary conditions. The existence of eigenvalues and oscillatory behavior of the associated eigenfunctions is established. The results extend those of G. J. Etgen and the author which were concerned with the homogeneous problem.

1. Introduction. The study of second order differential systems involving a parameter together with various kinds of boundary conditions has been an important part of mathematics for the past one hundred years, and subsequently the literature on this subject is voluminous. The purpose of this paper is to extend some results for these problems to differential systems which involve a "forcing term".

Consider the differential system

$$\text{(NH)} \qquad \begin{aligned} y' &= k(x, \lambda)z, \\ z' &= g(x, \lambda)y + f(x, \lambda), \end{aligned}$$

together with the associated homogeneous system

$$\text{(H)} \qquad \begin{aligned} u' &= k(x, \lambda)v, \\ v' &= g(x, \lambda)u, \end{aligned}$$

where $k(x, \lambda)$, $g(x, \lambda)$, and $f(x, \lambda)$ are real-valued functions on

$$X: a \leqq x \leqq b, \quad L: \lambda_{\#} - \delta < \lambda < \lambda_{\#} + \delta, \qquad 0 < \delta \leqq \infty, \quad -\infty < a < b < \infty.$$

The system (NH) shall be considered together with two-point boundary conditions of the form

$$\text{(1a)} \qquad \alpha(\lambda)y(a, \lambda) - \beta(\lambda)z(a, \lambda) = 0,$$

$$\text{(1b)} \qquad \gamma_1(\lambda)y(a, \lambda) + \delta_1(\lambda)z(a, \lambda) = \gamma_2(\lambda)y(b, \lambda) + \delta_2(\lambda)z(b, \lambda) + H(b, \lambda),$$

where $H(x, \lambda) = \int_a^x h(t, \lambda)y(t, \lambda)\,dt$, and $\alpha, \beta, \gamma_i, \delta_i, i = 1, 2$, and h are real-valued functions on L and XL, respectively.

The problem (H), (1a, b) has been studied by G. J. Etgen and the author [2] and by C. Comstock and P. Dunne [1]. Each of these papers is an extension of the fundamental work of W. M. Whyburn [11]. Nonhomogeneous boundary problems without integral terms have been studied by the author [7], [8]. This work will extend that of the above-named authors.

Assuming that $k(x, \lambda)$ is nonzero in XL, the system (NH) may be written as

$$\text{(2)} \qquad (y'/k)' - gy = f$$

and the associated homogeneous equation is

$$\text{(3)} \qquad (u'/k)' - gu = 0.$$

---

* Received by the editors February 26, 1976, and in final revised form October 26, 1976.

† Department of Mathematics, University of Alabama in Birmingham, Birmingham, Alabama. Now at Exxon Production Research Company, Houston, Texas 77001.

It is easily verified that the general solution of (2) is given by

$$(4)\quad y(x,\lambda)=\left[c_1(\lambda)-\int_a^x f(t,\lambda)v(t,\lambda)\,dt\right]u(x,\lambda)+\left[c_2(\lambda)+\int_a^x f(t,\lambda)u(t,\lambda)\,dt\right]v(x,\lambda),$$

where $\{u(x,\lambda), v(x,\lambda)\}$ is a solution basis of (3) such that

$$(uv'-vu')/k\equiv 1 \quad \text{on } XL.$$

Substituting (4) into the boundary conditions, one obtains

$$c_1B_1(u)+c_2B_1(v)=0,$$
$$c_1B_2(u)+c_2B_2(v)=-B_2(y_p),$$

where $B_1(u)=\alpha(\lambda)u(a,\lambda)-\beta(\lambda)u'(a,\lambda)/k(a,\lambda)$, $B_2(u)=\gamma_1(\lambda)u(a,\lambda)+\delta_1(\lambda)\cdot(u'(a,\lambda)/k(a,\lambda))-\gamma_2(\lambda)u(b,\lambda)-\delta_2(\lambda)(u'(b,\lambda)/k(b,\lambda))-\int_a^b h(t,\lambda)u(t,\lambda)\,dt$, and $y_p(x,\lambda)$ is obtained from (4) by putting $c_1=c_2=0$. (Note that $y_p(a)=y_p'(a)=0$.) Hence, as is the case for other nonhomogeneous boundary problems, the problem (NH), (1a, b) has a unique solution for those values of λ for which the associated homogeneous problem has only the trivial solution. Further, for those values of λ for which the problem (H), (1a, b) has a nontrivial solution, the nonhomogeneous problem either has no solution or infinitely many solutions.

Max Mason [6] considered a nonhomogeneous equation of the form (2) together with self-adjoint boundary conditions and obtained necessary and sufficient conditions for existence of eigenvalues when the homogeneous problem had a nontrivial solution. For the problem (NH), (1a, b), these conditions reduce to the condition that $B_2(y_p)=0$.

In either case, the homogeneous problem must be solved or shown to have only the trivial solution before solutions of the nonhomogeneous problem can be determined. This seems to be an indirect and highly ineffective method of solution.

The following hypotheses on the coefficients involved in the boundary problem will be assumed throughout:

(H_1) For each $x\in X$, each of $k(x,\lambda)$, $g(x,\lambda)$, $h(x,\lambda)$, and $f(x,\lambda)$ is continuous on L.

(H_2) For each $\lambda\in L$, each of $k(x,\lambda)$, $g(x,\lambda)$, $h(x,\lambda)$, and $f(x,\lambda)$ is measurable on X.

(H_3) There exists a Lebesgue integrable function $M(x)$ on X such that $|k(x,\lambda)|\leqq M(x)$, $|g(x,\lambda)|\leqq M(x)$, $|h(x,\lambda)|\leqq M(x)$, and $|f(x,\lambda)|\leqq M(x)$ on XL.

(H_4) $k(x,\lambda)>0$ on XL.

(H_5) Each of the functions $\alpha(\lambda)$, $\beta(\lambda)$, $\gamma_i(\lambda)$, and $\delta_i(\lambda)$, $i=1, 2$, is continuous on L.

(H_6) $\alpha^2(\lambda)+\beta^2(\lambda)>0$ on L.

(H_7) $\beta(\lambda)\geqq 0$ on L.

(H_8) $\gamma_i^2(\lambda)+\delta_i^2(\lambda)>0$ on L, $i=1, 2$.

(H_9) $g(x,\lambda)$ is not identically zero on any subinterval of X for each $\lambda\in L$ and is not identically zero on any subinterval of L for any $x\in X$.

2. Preliminary results. We seek to establish the existence of values of $\lambda\in L$ for which there corresponds a nontrivial solution of (NH) satisfying (1a, b). Such values of λ are called *eigenvalues* of the boundary problem. By a *nontrivial solution* of (NH), we mean a solution pair $\{y(x,\lambda), z(x,\lambda)\}$ of (NH) such that $y^2(x,\lambda)+z^2(x,\lambda)\not\equiv 0$ on XL.

As in [9], it can be shown that if hypotheses (H_1)–(H_3) are satisfied and if $\{y(x,\lambda), z(x,\lambda)\}$ is any solution pair of (NH) such that $y^2(x,\lambda)+z^2(x,\lambda)>0$ on XL, then there exists a pair of functions $\rho(x, \lambda)$, $\theta(x, \lambda)$ with the property that

$$(5)\qquad \begin{aligned} y(x, \lambda) &= \rho(x, \lambda) \sin \theta(x, \lambda), \\ z(x, \lambda) &= \rho(x, \lambda) \cos \theta(x, \lambda). \end{aligned}$$

Furthermore, $\rho(x, \lambda)$ and $\theta(x, \lambda)$ satisfy the differential equations

$$(6)\qquad \begin{aligned} &\text{(a)}\quad \rho' = \rho(k+g)(\sin\theta)\cos\theta + f\cos\theta, \\ &\text{(b)}\quad \theta' = k\cos^2\theta - g\sin^2\theta - (f\sin\theta)/\rho, \end{aligned}$$

with the initial conditions $\rho(a) > 0$; $0 \leqq \theta(a) \leqq 2\pi$.

It is required that $y^2(x, \lambda) + z^2(x, \lambda) > 0$ on XL to insure that $\rho'(x, \lambda)$ and $\theta'(x, \lambda)$ are defined on XL. The following theorem provides conditions under which these functions are well defined [8].

THEOREM 1. *Suppose that* $\rho(a, \lambda) + 1 > \exp \int_a^b M(s)\, ds$ *on* L, *where* $\rho(x, \lambda)$ *is defined by* (6), *and* $M(x)$ *is the Lebesgue integrable bound of the coefficients. Then* $\rho(x, \lambda) > 0$ *and consequently*, $y^2(x, \lambda) + z^2(x, \lambda) > 0$ *on* XL.

Hypotheses (H_1)–(H_3) allow application of fundamental existence theorems for differential systems to obtain the existence of a solution pair $\{y(x, \lambda), z(x, \lambda)\}$ of (NH) on XL such that

$$(7)\qquad y(a, \lambda) \equiv \beta(\lambda); \quad z(a, \lambda) \equiv \alpha(\lambda)$$

on L. Further, $\{y(x, \lambda), z(x, \lambda)\}$ has the polar coordinate representation (5), with $\{\rho(x, \lambda), \theta(x, \lambda)\}$ solutions of (6) with the initial conditions

$$(7')\qquad \begin{aligned} &\rho(a, \lambda) = [\alpha^2(\lambda) + \beta^2(\lambda)]^{1/2}; \qquad \sin\theta(a, \lambda) \equiv \beta(\lambda)/[\alpha^2(\lambda) + \beta^2(\lambda)]^{1/2}, \\ &\cos\theta(a, \lambda) \equiv \alpha(\lambda)/[\alpha^2(\lambda) + \beta^2(\lambda)]^{1/2}, \qquad 0 \leqq \theta(a, \lambda) \leqq 2\pi. \end{aligned}$$

Since $\beta(\lambda) \geqq 0$, we may assume $0 \leqq \theta(a, \lambda) \leqq \pi$.

Using this nonhomogeneous polar coordinate transformation, we may now state an existence theorem for eigenvalues of (NH), (1a, b).

THEOREM 2. *Let* $\{\rho(x, \lambda), \theta(x, \lambda)\}$ *be the solution pair of* (6) *defined by* (7′) *and let* $\{y(x, \lambda), z(x, \lambda)\}$ *be defined by* (5). *Assume the hypothesis of Theorem* 1. *Then* $\{y(x, \lambda), z(x, \lambda)\}$, *satisfies* (NH), (1a) *and* $y^2 + z^2 > 0$ *on* XL. *Further*, $\theta(b, \lambda) \geqq 0$ *on* L. *In addition to* (H_1)–(H_9), *let the following conditions hold*:

(i) $h(x, \lambda)/g(x, \lambda)$ *is defined, integrable, nonnegative and nondecreasing on* X *for each* $\lambda \in L$;

(ii) $\rho(b, \lambda) \geqq \rho(x, \lambda)$ *on* X *for each* $\lambda \in L$;

(iii) $\int_a^b [h(t, \lambda)f(t, \lambda)/g(t, \lambda)]\, dt$ *is of one sign on* L;

(iv) *either*

$$\rho(b, \lambda)\delta_2(\lambda) \geqq [\gamma_1^2(\lambda) + \delta_1^2(\lambda)]^{1/2} + \left| \int_a^b [h(t, \lambda)f(t, \lambda)/g(t, \lambda)]\, dt \right|,$$

for each $\lambda \in L$, *or*

$$\rho(b, \lambda)[\delta_2(\lambda) + 2h(b, \lambda)/g(b, \lambda)] \leqq -[\gamma_1^2(\lambda) + \delta_1^2(\lambda)]^{1/2} - \left| \int_a^b [h(t, \lambda)f(t, \lambda)/g(t, \lambda)]\, dt \right|,$$

for each $\lambda \in L$.

If m *is the least nonnegative integer such that* $\inf \theta(b, \lambda) < m\pi$ *and* n *is an integer such that* $\sup \theta(b, \lambda) > n\pi$, *and if* $n \geqq m+1$, *then there exist at least* p, $p = n - m$, *nonempty sets of eigenvalues* $T_0, T_1, \cdots, T_{p-1}$ *for the boundary problem* (NH), (1a, b).

Furthermore, the number of distinct eigenvalues for (NH), (1a, b) *is at least $p/2$ if p is even and at least $(p+1)/2$ if p is odd.*

Proof. Since many of the techniques used in this proof are similar to those used in [2], the reader is referred to [2] for details.

If $\{\rho(x,\lambda), \theta(x,\lambda)\}$ is the solution pair of (6) defined by (7′) and $\{y(x,\lambda), z(x,\lambda)\}$ is defined by (5), then $\{y, z\}$ is a solution of (NH) satisfying (7). Boundary condition (1a) is clearly satisfied by this solution pair.

The coefficient hypotheses imply that $\theta(x,\lambda)$ is continuous on XL. Since $\theta(a,\lambda)\geqq 0$ on L and since $\theta'(x,\lambda)>0$ when $y(x,\lambda)=0$, for any $\lambda\in L$, then $\theta(b,\lambda)\geqq 0$ on L.

Since $\theta(b,\lambda)$ ranges in value from less than $m\pi$ to more than $n\pi$, there exist values of λ, λ_0 and λ_p, such that $\theta(b,\lambda_0)=m\pi$ and $\theta(b,\lambda_p)=n\pi$, and we can assume without loss of generality that $\lambda_0<\lambda_p$.

From (5), boundary condition (1b) becomes

$$(8)\qquad [\gamma_1^2(\lambda)+\delta_1^2(\lambda)]^{1/2}\sin[\theta(a,\lambda)+\tau(\lambda)]=\rho(b,\lambda)G(b,\lambda)+H(b,\lambda),$$

where

$$(9)\qquad \begin{gathered}\sin\tau=\delta_1/(\gamma_1^2+\delta_1^2)^{1/2},\qquad \cos\tau=\gamma_1/(\gamma_1^2+\delta_1^2)^{1/2},\\ G(x,\lambda)=\gamma_2(\lambda)\sin\theta(x,\lambda)+\delta_2(\lambda)\cos\theta(x,\lambda).\end{gathered}$$

Define $Q(\lambda)$ by

$$(10)\qquad Q(\lambda)=\rho(b,\lambda)G(b,\lambda)+H(b,\lambda).$$

Fix $\lambda\in L$ and consider

$$\begin{aligned}H(b,\lambda)&=\int_a^b h(t,\lambda)y(t,\lambda)\,dt=\int_a^b h(t,\lambda)[(z'(t,\lambda)-f(t,\lambda))/g(t,\lambda)]\,dt\\ &=\int_a^b [h(t,\lambda)z'(t,\lambda)/g(t,\lambda)]\,dt-m(\lambda),\end{aligned}$$

where

$$(11)\qquad m(\lambda)=\int_a^b [h(t,\lambda)f(t,\lambda)/g(t,\lambda)]\,dt.$$

Hypothesis (H_9) together with condition (i) allow application of a mean value theorem for integrals [5, Thm. 244, p. 164] to obtain

$$(12)\qquad \frac{h(b,\lambda)}{g(b,\lambda)}\int_{\bar{x}}^b z'(t,\lambda)\,dt\leqq H(b,\lambda)+m(\lambda)\leqq\frac{h(b,\lambda)}{g(b,\lambda)}\int_{x^*}^b z'(t,\lambda)\,dt,$$

where $\bar{x}$ and x^* are such that

$$\min_{x\in X}\int_x^b z'(t,\lambda)\,dt=\int_{\bar{x}}^b z'(t,\lambda)\,dt\quad\text{and}\quad \max_{x\in X}\int_x^b z'(t,\lambda)\,dt=\int_{x^*}^b z'(t,\lambda)\,dt.$$

Thus we have

$$(13)\qquad \begin{aligned}&\rho(b,\lambda)\left\{G(b,\lambda)+\frac{h(b,\lambda)}{g(b,\lambda)}\left[\cos\theta(b,\lambda)-\frac{\rho(\bar{x},\lambda)}{\rho(b,\lambda)}\cos\theta(\bar{x},\lambda)\right]\right\}-m(\lambda)\\ &\leqq Q(\lambda)\\ &\leqq\rho(b,\lambda)\left\{G(b,\lambda)+\frac{h(b,\lambda)}{g(b,\lambda)}\left[\cos\theta(b,\lambda)-\frac{\rho(x^*,\lambda)}{\rho(b,\lambda)}\cos\theta(x^*,\lambda)\right]\right\}-m(\lambda).\end{aligned}$$

Since $n=m+p$, $p \geqq 1$, and since $\theta(b, \lambda)$ is continuous in λ, there are $p-1$ values of λ, $\lambda_1<\lambda_2<\cdots<\lambda_{p-1}$ on (λ_0, λ_p) such that $\theta(b, \lambda)=(m+j)\pi$, $j=1, 2, \cdots, p-1$. Choose any integer j, $0 \leqq j \leqq p-1$, and assume $\cos \theta(b, \lambda_j)=\pm 1$. Then $\cos \theta(b, \lambda_{j+1})=-1$. From (13) and (ii), it follows that

$$\begin{aligned} &\Gamma(\lambda_j)-m(\lambda_j) \leqq Q(\lambda_j) \leqq \Delta(\lambda_j)-m(\lambda_j), \\ &-\Delta(\lambda_{j+1})-m(\lambda_{j+1}) \leqq Q(\lambda_{j+1}) \leqq -\Gamma(\lambda_{j+1})-m(\lambda_{j+1}), \end{aligned} \tag{14}$$

where

$$\begin{aligned} &\Gamma(\lambda)=\rho(b, \lambda)\delta_2(\lambda), \\ &\Delta(\lambda)=\rho(b, \lambda)[\delta_2(\lambda)+2h(b, \lambda)/g(b, \lambda)]. \end{aligned} \tag{15}$$

Suppose the first condition of (iv) holds on L. If $m(\lambda) \geqq 0$ on L, then $\Gamma(\lambda) \geqq (\gamma_1^2+\delta_1^2)^{1/2}+m(\lambda) \geqq (\gamma_1^2+\delta_1^2)^{1/2}-m(\lambda)$. Hence $(\gamma_1^2+\delta_1^2)^{1/2} \leqq \Gamma(\lambda)-m(\lambda)$ and $-\Gamma(\lambda)-m(\lambda) \leqq -(\gamma_1^2+\delta_1^2)^{1/2}$, so that $[\gamma_1^2(\lambda_j)+\delta_1^2(\lambda_j)]^{1/2} \leqq Q(\lambda_j)$ and $Q(\lambda_{j+1}) \leqq -[\gamma_1^2(\lambda_{j+1})+\delta_1^2(\lambda_{j+1})]^{1/2}$. A similar result holds if $m(\lambda) \leqq 0$ on L. Likewise, if (iii) and the second condition of (iv) is assumed, we have that as λ increases from λ_j to λ_{j+1}, $Q(\lambda)$ changes continuously in value from not less than $[\gamma_1^2(\lambda)+\delta_1^2(\lambda)]^{1/2}$ to not more than $-[\gamma_1^2(\lambda)+\delta_1^2(\lambda)]^{1/2}$ (or vice versa) for $j=0, 1, \cdots, p-1$.

We have thus established that the continuous curves $(\gamma_1^2+\delta_1^2)^{1/2} \sin[\theta(a, \lambda)+\tau(\lambda)]$ and $Q(\lambda)$ must intersect at least once on the intervals $[\lambda_j, \lambda_{j+1}]$, $j=0, 1, \cdots, p-1$. Hence the boundary condition (1b) is satisfied at least once on $[\lambda_j, \lambda_{j+1}]$. Let $T_j=\{\lambda \in [\lambda_j, \lambda_{j+1}] | \text{(1b) is satisfied}\}$, $j=0, 1, \cdots, p-1$. It could be the case that the λ which satisfy (1b) are alternate endpoints. Thus there will be at least $p/2$ or $(p+1)/2$ distinct eigenvalues for (NH), (1a, b), depending on whether p is even or odd. This completes the proof of the theorem.

Remarks. It should be noted that there may exist additional eigenvalues for (NH), (1a, b) outside of $[\lambda_0, \lambda_p]$, and that each of the nonempty sets of eigenvalues T_j, $j=0, 1, \cdots, p-1$ can be finite, countable or uncountable.

Conditions (ii) and (iv) of the hypothesis of the theorem concern the amplitudes of the solution pair $\{y(x, \lambda), z(x, \lambda)\}$. Solving equation (6a), one obtains

$$\rho(x, \lambda)=\exp w(x, \lambda)\left\{\rho(a, \lambda)+\int_a^x [f(t, \lambda) \cos \theta(t, \lambda)] \exp(-w(t, \lambda))\, dt\right\},$$

where

$$w(x, \lambda)=\int_a^x [k(t, \lambda)+g(t, \lambda)](\sin \theta(t, \lambda)) \cos \theta(t, \lambda)\, dt.$$

Since

$$\begin{aligned} -\int_x^b (f \cos \theta) \exp(-w)\, dt &\leqq \left|\int_x^b (f \cos \theta) \exp(-w)\, dt\right| \\ &\leqq \exp \int_a^b M(s)\, ds-\exp \int_a^x M(s)\, ds, \end{aligned}$$

it follows from the hypothesis of Theorem 1 that $\rho(a, \lambda)+\int_x^b (f \cos \theta) \exp(-w)\, dt$ is nonnegative on XL. If one assumes, for example, that $w(b, \lambda) \geqq w(x, \lambda)$ for $x \in X$, then condition (ii) will be satisfied.

Further, one should note that the first of conditions (iv) tacitly implies that $\delta_2(\lambda)>0$ on L, while the second implies $\delta_2(\lambda)<0$ on L. Also, in keeping with [1], [2], the first of conditions (iv) will be satisfied if one can verify that $w(b,\lambda)\geqq 0$ on L and that $\delta_2(\lambda)\geqq[\gamma_1^2(\lambda)+\delta_1^2(\lambda)]^{1/2}+|\int_a^b[h(t,\lambda)f(t,\lambda)/g(t,\lambda)]\,dt|$, on L.

The following corollary is stated for completeness. Since the proof is similar to that of [2, Cor. 2], it is omitted.

COROLLARY. *Under the hypothesis of the theorem, there exist p nonempty sets of eigenvalues $J_0,\cdots,J_{p-1}$ for the problem* (NH), (1a, b) *such that if $\rho_j\in J_j$, $j=0,1,\cdots,p-1$, then $\theta(b,\rho_j)\geqq(m+j-1)\pi$. Further, if $\rho_j\in J_j$, then the corresponding solution $\{y(x,\rho_j),z(x,\rho_j)\}$ has at least $m+j-1$ zeros on X, when j is such that $m+j-1\geqq 0$.*

It is clear that if boundary condition (1b) is replaced by the condition

$$\gamma_1(\lambda)y(a,\lambda)+\delta_1(\lambda)z(a,\lambda)=\gamma_2(\lambda)y(b,\lambda)+\delta_2(\lambda)z(b,\lambda)+J(b,\lambda), \tag{1c}$$

where $J(b,\lambda)=\int_a^b j(t,\lambda)z(t,\lambda)\,dt$, then a similar theorem may be proved. Of course, the $m(\lambda)$ term will not be involved in this result.

We now consider the system (NH) together with the boundary conditions

$$\alpha(\lambda)y(a,\lambda)-\beta(\lambda)z(a,\lambda)=0, \tag{16a}$$

$$\gamma_1(\lambda)y(a,\lambda)-\delta_1 z(a,\lambda)=\gamma_2(\lambda)y(b,\lambda)-\delta_2(\lambda)z(b,\lambda)+M(b,\lambda), \tag{16b}$$

where $M(b,\lambda)=\int_a^b p(t,\lambda)[\sigma(\lambda)y(t,\lambda)-\xi(\lambda)z(t,\lambda)]\,dt$, and $\sigma(\lambda)$ and $\xi(\lambda)$ are continuous functions on L, and that $\gamma_2(\lambda)\xi(\lambda)-\delta_2(\lambda)\sigma(\lambda)\neq 0$ on L. Define the functions $s(x,\lambda)$ and $t(x,\lambda)$ by

$$\begin{aligned} s(x,\lambda)&=\gamma_2(\lambda)y(x,\lambda)-\delta_2(\lambda)z(x,\lambda),\\ t(x,\lambda)&=\sigma(\lambda)y(x,\lambda)-\xi(\lambda)z(x,\lambda). \end{aligned} \tag{17}$$

Then,

$$\begin{aligned} s'&=[(\gamma_2\sigma-\delta_2\xi)s+(\delta_2^2g-\gamma_2^2k)t-\delta_2(\gamma_2\xi-\delta_2\sigma)f]/(\gamma_2\xi-\delta_2\sigma),\\ t'&=[(\sigma^2k-\xi^2g)s+(\delta_2\xi g-\gamma_2\sigma k)t-\xi(\gamma_2\xi-\delta_2\sigma)f]/(\gamma_2\xi-\delta_2\sigma). \end{aligned} \tag{18}$$

Assume, without loss of generality, that $\gamma_2\xi-\delta_2\sigma\equiv 1$ on L and define S, T, K, G, U, V, and ϕ by

$$\begin{aligned} &S=s\,e^{-\phi},\quad T=t\,e^{\phi},\quad K=(\delta_2^2g-\gamma_2^2k)\,e^{-2\phi},\\ &G=(\sigma^2k-\xi^2g)\,e^{2\phi},\qquad \phi=\int_a^x(\gamma_2\sigma k-\delta_2\xi g)\,dt,\\ &U=\delta_2 f\,e^{-\phi},\qquad V=\xi f\,e^{\phi}. \end{aligned} \tag{19}$$

Using these functions, the boundary problem (NH), (16a, b) is expressed as

$$\begin{aligned} S'&=KT-U,\\ T'&=GS-V, \end{aligned} \tag{20}$$

$$A(\lambda)S(a,\lambda)-B(\lambda)T(a,\lambda)=0, \tag{21a}$$

$$C(\lambda)S(a,\lambda)-D(\lambda)T(a,\lambda)=E(\lambda)S(b,\lambda)+N(b,\lambda), \tag{21b}$$

where

$$(22)\qquad \begin{aligned} &A=\alpha\xi-\beta\sigma, \quad B=\gamma_2\beta-\delta_2\alpha, \quad C=\gamma_1\xi-\delta_1\sigma,\\ &D=\gamma_2\delta_1-\gamma_1\delta_2, \qquad E=\exp\phi(b,\lambda),\\ &N(b,\lambda)=\int_a^b p(u,\lambda)\exp(-\phi(u,\lambda))T(u,\lambda)\,du. \end{aligned}$$

It is clear that by using the techniques developed to solve the problem (NH), (1a, b), we can obtain the existence of eigenvalues for the problem (20), (21). The appearance of the additional forcing function $U(x,\lambda)$ presents no difficulties since a polar coordinate transformation can be made for the system (20) [9]. Thus, by imposing conditions on the coefficients of (20), (21) which are analogous to those of (NH), (1a, b), the existence of eigenvalues may be established. By reversing the transformations defined by (17) and (19), the problem (NH), (16a, b) may be solved.

For closing remarks, we consider the question of being able to find integers m and n having the properties described in the hypothesis of Theorem 2. Since $k(x,\lambda)>0$ on XL, we consider the equation (2).

Let $u(x,\lambda)$ and $v(x,\lambda)$ be linearly independent solutions of (3) satisfying the initial conditions

$$\begin{aligned} &u(a,\lambda)\equiv 1, && v(a,\lambda)\equiv 0,\\ &u'(a,\lambda)/k(a,\lambda)\equiv 0, && v'(a,\lambda)/k(a,\lambda)\equiv 1, \end{aligned}$$

on L. Then $(uv'-vu')/k\equiv 1$ on L and the general solution of (2) is given by (4). If we take $y(x,\lambda)$ to be the particular solution of (2) satisfying the initial conditions $y(a,\lambda)\equiv\beta(\lambda)$, $y'(a,\lambda)/k(a,\lambda)\equiv\alpha(\lambda)$, then $c_1(\lambda)\equiv\beta(\lambda)$ and $c_2(\lambda)\equiv\alpha(\lambda)$. Further,

$$(23)\qquad [u(x,\lambda)y'(x,\lambda)-u'(x,\lambda)y(x,\lambda)]/k(x,\lambda)=\alpha(\lambda)+\int_a^x f(t,\lambda)u(t,\lambda)\,dt,$$

and

$$(24)\qquad [y(x,\lambda)v'(x,\lambda)-v(x,\lambda)y'(x,\lambda)]/k(x,\lambda)=\beta(\lambda)-\int_a^x f(t,\lambda)v(t,\lambda)\,dt.$$

Hence, if either $\alpha(\lambda)+\int_a^b f(t,\lambda)u(t,\lambda)\,dt$ or $\beta(\lambda)-\int_a^b f(t,\lambda)v(t,\lambda)\,dt$ is nonzero on L, then the zeros of y and v or y and u separate each other on L. Thus, for example, if $\beta(\lambda)+1>\exp\int_a^b M(t)\,dt$ on L, then the zeros of y and v will separate on L. (Here $M(x)$ is the Lebesgue integrable bound of the functions k, g, and f.) Now by applying oscillation criteria to solutions of (H), we obtain oscillation of solutions of (NH) on L. Since the oscillation of $y(b,\lambda)$ on L is equivalent to $\theta(b,\lambda)$ going from less than $m\pi$ to more than $n\pi$, for integers m and n, the hypotheses of Theorem 2 may be verified. Conditions guaranteeing oscillation of solutions of (H) on L may be found in Ince [4, pp. 231–237], Whyburn [10, p. 852], or Ettlinger [3, pp. 136, 137].

REFERENCES

[1] C. Comstock and P. W. Dunne, *Second order equations with two point and integral boundary conditions*, Notices Amer. Math. Soc., 20 (1973), A-523; submitted for publication.

[2] G. J. Etgen and S. C. Tefteller, *Second order differential equations with general boundary conditions*, this Journal, 3 (1972), pp. 512–519.

[3] H. J. ETTLINGER, *Oscillation theorems for the real, self-adjoint linear system of the second order*, Trans. Amer. Math. Soc., 22 (1921), pp. 136–143.
[4] E. L. INCE, *Ordinary Differential Equations*, Dover, New York, 1956.
[5] H. KESTELMAN, *Modern Theories of Integration*, Dover, New York, 1960.
[6] MAX MASON, *On the boundary value problems of linear ordinary differential equations of second order*, Trans. Amer. Math. Soc., 7 (1906), pp. 337–360.
[7] S. C. TEFTELLER, *A two-point boundary problem for nonhomogeneous second order differential equations*, Pacific J. Math., 53 (1974), pp. 635–642.
[8] ———, *Boundary value problems for second order nonhomogeneous differential systems*, Proc. Amer. Math. Soc., 52 (1975), pp. 271–278.
[9] ———, *A polar coordinate transformation for nonhomogeneous differential systems*, Portugal. Math., 34 (1975), pp. 219–223.
[10] W. M. WHYBURN, *Existence and oscillation theorems for non-linear differential systems of the second order*, Trans. Amer. Math. Soc., to appear.
[11] ———, *Second-order differential systems with integral and k-point boundary conditions*, Ibid., 30 (1928), pp. 630–640.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0002 $01.00/0

DUAL ORTHOGONAL SERIES WITH OSCILLATORY MODIFIERS*

ROBERT P. FEINERMAN† AND ROBERT B. KELMAN‡

Abstract. Dual orthogonal series with oscillatory modifiers occur in problems of communication theory. We consider such a problem in an abstract Hilbert space and prove theorems of existence and uniqueness for dual series in which the ratios of the modifiers oscillate between finite, nonnegative limits. The analysis is based upon properties of appropriately constructed linear functionals.

1. Introduction. The mixed boundary value problems of mathematical physics have provided the main impetus for studying dual orthogonal series ([1], [2], [3]; cf. [4] for an exposition not based on applications). In considering the enciphering of a message we have been led to a dual orthogonal series problem somewhat different from those studied earlier. (In what follows, we use the notation and definitions given in [2, § 2].) Namely, in dual orthogonal series associated with mixed boundary value problems, one inevitably finds that the ratio of the modifiers, b_n/a_n, tends to a nonnegative limit, possibly infinite, as n tends to infinity, whereas in cryptological problems b_n/a_n may oscillate as n tends to infinity. It is series with this latter property that are studied here.

As an idealized application consider a message $f(t)$ of T seconds duration encoded using a sequence $\{\phi_n(t)\}$ which is orthonormal on $0<t<T$ (vid. [5, esp. Chap. 2] and [6, esp. Chap. 13] for general background). For simplicity let us assume the absence of noise. The transmitted message encoded without disguise is $\sum f_n\phi_n(t)$ where $f_n = \int_0^T f\phi_n\,dt$. To encipher the message we choose a sequence of nonzero constants $\{a_n\}$ and transmit $\sum f_n a_n \phi_n(t)$. A decoder can be constructed so that given $\{a_n\}$ the message is resynthesized. Consider now a procedure for more effectively diguising the message. Partition T into disjoint measurable subsets T_1 and T_2, and let $\{b_n\}$ be a second sequence of constants. Let $\sum f_n a_n \phi_n(t)$ be transmitted for $t \in T_1$ and $\sum f_n b_n \phi_n(t)$ be transmitted for $t \in T_2$. Then the first theorem below states that a decoder which deciphers the message is physically realizable provided that b_n/a_n $(n = 1, 2, \cdots)$ is bounded above zero and below infinity. In practice if a_n and b_n are positive, this is always achieved, since only a finite number of the functions $\{\phi_n(t)\}$ are used in the transmission of information. The second theorem weakens the restriction on the modifiers by allowing a finite number of them to be zero, which, as explained in [2] and [10], may occur in applications. Although Theorem 2 is essentially a technical extension of Theorem 1, its proof is more involved even assuming the results of Theorem 1 and [7]. Roughly speaking, this corresponds to the extra effort required to show completeness of Sturm–Liouville eigenfunctions in the presence of a zero eigenvalue [9, p. 246].

The theorems given here generalize and complement our earlier results [2], [7], [10] illustrating how a suitable application can aid the development of a purely mathematical theory.

2. Existence theorems. We denote by **R** a real, separable abstract Hilbert space and by ℓ^2 the real Hilbert space of square summable column vectors. Let $\{a_n\}$ and $\{b_n\}$

* Received by the editors February 26, 1976. It is noted with appreciation that the first version of this paper was written while the second author was a guest of The Hebrew University of Jerusalem. The research of the second author was supported in part by the U.S. Army Research Office under Grant DAH C04 75 G0162.

† Department of Mathematics, H. H. Lehman College, City University of New York, Bronx, New York 10468.

‡ Department of Computer Science, Colorado State University, Fort Collins, Colorado 80523.

be sequences of nonnegative constants. The subspaces **P** and **Q** of **R** are orthogonal complements. P and Q denote respectively the projection operators from **R** onto **P** and **Q**. Let ψ_n be defined by $\psi_n = a_n P\phi_n + b_n Q\phi_n$ where $\{\phi_n\}$ is a complete, orthonormal sequence in **R**. Our goal is to prove certain existence and uniqueness theorems for the solution j, a real infinite vector not necessarily in ℓ^2, to the dual orthogonal series equation

$$\sum_{n=1}^{\infty} j_n \psi_n = f \tag{1}$$

given $\{a_n\}$, $\{b_n\}$, and $f \in \mathbf{R}$. We start with the following lemma.

LEMMA 1. *Let $\{\varepsilon_n\}$ be a sequence of constants bounded below two and above zero. Let $\xi_n = P\phi_n + \varepsilon_n Q\phi_n$. Then the equation*

$$\sum_{n=1}^{\infty} j_n \xi_n = f \tag{2}$$

has a unique solution j and $j \in \ell^2$.

Proof. Let T be a linear operator from **R** to **R** defined by $T\phi_n = \xi_n$. Note that

$$T\phi_n = \phi_n + (\varepsilon_n - 1)Q\phi_n.$$

Since $\{\phi_n\}$ is a maximal orthonormal set, the domain of T is **R**. Obviously T is symmetric so that by the Hellinger–Toeplitz theorem [8, p. 51] it is bounded. On the other hand (using the notation $f_n = (f, \phi_n)$) we see that

$$\|Tf\| \geqq \left\| \sum_{n=1}^{\infty} f_n \phi_n \right\| - \left\| \sum_{n=1}^{\infty} f_n (1-\varepsilon_n)\phi_n \right\| \geqq \varepsilon \|f\|$$

where

$$\varepsilon = 1 - \sup\{|1-\varepsilon_n| : n = 1, 2, \cdots\}.$$

Since ε is positive, $\lambda = 0$ is a regular point for T. Thus T^{-1} exists as an everywhere defined bounded operator on **R** [8, p. 95]. If we seek j such that

$$\sum_{n=1}^{\infty} j_n \phi_n = T^{-1} f, \tag{3}$$

then j is the unique element of ℓ^2 given by $j_n = (T^{-1}f, \phi_n)$. Now j is a solution to (3) if and only if it is a solution to (2) which completes the proof of the lemma.

From this result we can easily establish

THEOREM 1. *Let $\{a_n\}$ and $\{b_n\}$ be sequences of positive constants such that the ratio b_n/a_n $(n = 1, 2, \cdots)$ is bounded above zero and below infinity. Then* (1) *has a unique solution j and $\sum (j_n a_n)^2 < \infty$. Therefore if $\{a_n : n = 1, 2, \cdots\}$ is bounded above zero, $j \in \ell^2$.*

Proof. From the hypothesis we know that there are positive constants μ and M such that

$$\mu < \frac{b_n}{M a_n} < 2 - \mu, \qquad n = 1, 2, \cdots.$$

Define g, k_n and ξ_n by

$$g = Pf + \frac{Q}{M} f, \quad k_n = a_n j_n, \quad \xi_n = P\phi_n + \left(\frac{b_n}{M a_n}\right) Q\phi_n.$$

Then $\mathscr{j}$ is a solution to eq. (1) if and only if $\mathscr{k}$ is a solution to $\sum \mathscr{k}_n\xi_n = g$. Since $\{\xi_n\}$ satisfies the hypothesis of Lemma 1, the proof of Theorem 1 is complete.

Next we consider the case in which some of the modifiers are zero and start with a lemma for which it is convenient to use the notion of a basis, viz., a sequence of elements $\{\theta_n\}$ in $\mathbf{R}$ is a basis if $f \in \mathbf{R}$ can be represented in only one way as a series $f = \sum \mathscr{k}_n\theta_n$ where $\{\mathscr{k}_n\}$ is a sequence of real numbers.

LEMMA 2. *Let* $\{\theta_1, \theta_2, \cdots\}$ *be a basis and* $\{\gamma_1, \gamma_2, \cdots, \gamma_K\}$ *a finite sequence of elements in* $\mathbf{R}$. *Then the set of elements*

$$\Delta = \{\gamma_1, \gamma_2, \cdots, \gamma_K; \theta_{K+1}, \theta_{K+2}, \cdots\}$$

is a basis if and only if it is complete.

Proof. Clearly, if Δ is a basis it is complete. Let us assume Δ is complete and show that this implies Δ is a basis. Let L_i $(i = 1, 2, \cdots)$ be a functional defined on $\mathbf{R}$ by $L_i\theta_j = \delta_{ij}$ (Kronecker delta). Note that if $f = \sum \mathscr{k}_n\theta_n$, then $f = \sum (L_nf)\theta_n$.

Let us show that the elements $\{\gamma_1, \gamma_2, \cdots, \gamma_K\}$ are linearly independent. Assume the contrary. Then there are constants d_n $(n = 1, 2, \cdots, K)$ not all zero and such that $\sum_1^K d_j\gamma_j = 0$. Consequently, $\sum_1^K d_j(L_i\gamma_j) = 0$ $(i = 1, 2, \cdots, K)$ so that the $K \times K$ matrix D, where $D_{ij} = L_i\gamma_j$, is singular. Consequently, there is a nonzero column vector $c = (c_1, c_2, \cdots, c_K)$ such that $D^Tc = 0$ where D^T is the transpose of D. Also, $(\sum_1^K c_jL_j)\theta_i = 0$ $(i = K+1, K+2, \cdots)$. Since $\sum_1^K c_jL_j$ is a bounded linear functional the Riesz representation theorem [8, p. 36] implies there is $\chi \in \mathbf{R}$ such that

$$\left(\sum_{j=1}^{K} c_jL_j\right)f = (f, \chi)$$

for all $f \in \mathbf{R}$. Consequently $(\chi, \gamma_j) = 0$ $(j = 1, 2, \cdots, K)$ and $(\chi, \theta_j) = 0$ $(j = K+1, K+2, \cdots)$. Since Δ is complete, $\chi = 0$. This contradicts $c \neq 0$. Thus the elements $\{\gamma_1, \gamma_2, \cdots, \gamma_k\}$ are linearly independent.

For a given $f \in \mathbf{R}$, define the constants $\{\mathscr{k}_n\}$ by

$$\sum_{n=1}^{K} (L_i\gamma_n)\mathscr{k}_n = L_if, \qquad i = 1, 2, \cdots, K, \tag{4}$$

$$\mathscr{k}_i = L_if - \sum_{n=1}^{K} (L_i\gamma_n)\mathscr{k}_n, \qquad i = K+1, K+2, \cdots. \tag{5}$$

Now $\mathscr{k}_1, \mathscr{k}_2, \cdots, \mathscr{k}_K$ are uniquely determined by (4) because D is nonsingular. The remaining coefficients $\mathscr{k}_{K+1}, \mathscr{k}_{K+2}, \cdots$ are then uniquely determined by (5). Next let us show that

$$f = \sum_{n=1}^{K} \mathscr{k}_n\gamma_n + \sum_{n=K+1}^{\infty} \mathscr{k}_n\theta_n. \tag{6}$$

Using (5) we see that

$$f = \sum_{n=K+1}^{\infty} \mathscr{k}_n\theta_n + \sum_{j=1}^{K} (L_jf)\theta_j + \sum_{n=1}^{K} \mathscr{k}_n \sum_{j=K+1}^{\infty} (L_j\gamma_n)\theta_j. \tag{7}$$

Remembering that $\gamma_n = \sum (L_j\gamma_n)\theta_j$ and using (4), we get

$$\sum_{n=1}^{K} \mathscr{k}_n\gamma_n = \sum_{j=1}^{K} (L_jf)\theta_j + \sum_{n=1}^{K} \mathscr{k}_n \sum_{j=K+1}^{\infty} (L_j\gamma_n)\theta_j.$$

This equation used with (7) establishes (6).

It remains to show that $\mathscr{k}$ is unique. If it is not unique, then there is $\mathscr{m}$ such that

$$\sum_{n=1}^{K} \mathscr{m}_n \gamma_n + \sum_{n=K+1}^{\infty} \mathscr{m}_n \theta_n = 0. \tag{8}$$

This leads to

$$\sum_{n=1}^{K} (L_i \gamma_n) \mathscr{m}_n = - \sum_{n=K+1}^{\infty} (L_i \theta_n) \mathscr{m}_n, \qquad i = 1, 2, \cdots, K. \tag{9}$$

Clearly, the right hand side of (9) is zero. Since D is nonsingular, this implies $\mathscr{m}_n = 0\ (n = 1, 2, \cdots, K)$. In view of this and the fact that $\{\theta_1, \theta_2, \cdots\}$ is a basis, equation (8) implies $\mathscr{m}_n = 0\ (n = K+1, K+2, \cdots)$, which completes the proof.

Assume the elements $\{\psi_n : n = 1, 2, \cdots, K\}$ are linearly independent and that $a_n = 0\ (n = 1, 2, \cdots, K)$. Then the $K \times K$ Gramian determinant $|(Q\phi_i, \phi_j)|$ is nonzero [8, p. 13]. If, further, b_n/a_n is positive for $n = K+1, K+2, \cdots$, then $\{\psi_n : n = 1, 2, \cdots\}$ is complete by the main result in [7]. If, in addition, we assume b_n/a_n is bounded above zero and below infinity for $n = K+1, K+2, \cdots$, then Lemma 1 implies that $\{\phi_1, \phi_2, \cdots, \phi_K, \psi_{K+1}, \psi_{K+2}, \cdots\}$ is a basis. Therefore, Lemma 2 implies $\{\psi_n : n = 1, 2, \cdots\}$ is a basis. Thus (1) has a unique solution $\mathscr{j}$. Rewriting (1) as

$$\sum_{n=K+1}^{\infty} \mathscr{j}_n \psi_n = f - \sum_{n=1}^{K} \mathscr{j}_n b_n Q \phi_n,$$

we see that Theorem 1 implies $\sum_{K+1}^{\infty} (\mathscr{j}_n a_n)^2 < \infty$. Thus we have established

THEOREM 2. *Let K be a positive integer, and suppose that $a_n = 0\ (n = 1, 2, \cdots, K)$ and the elements $\{\psi_1, \psi_2, \cdots, \psi_K\}$ are linearly independent. If $b_n/a_n\ (n = K+1, K+2, \cdots)$ is bounded above zero and below infinity, then (1) has a unique solution $\mathscr{j}$ and $\sum_{K+1}^{\infty} (\mathscr{j}_n a_n)^2 < \infty$.*

REFERENCES

[1] I. N. SNEDDON, *Mixed Boundary Value Problems in Potential Theory*, North-Holland, Amsterdam, 1966.

[2] R. B. KELMAN AND R. P. FEINERMAN, *Dual orthogonal series*, this Journal, 5 (1974), pp. 489–502.

[3] R. P. FEINERMAN, R. B. KELMAN AND C. A. KOPER, JR., *Dual orthogonal series: A case study of the influence of computing upon mathematical theory*, Proc. Sympos. Appl. Math., 20 (1974), pp. 129–134.

[4] I. C. GOHBERG AND I. A. FEL'DMAN, *Convolution Equations and Projection Methods for Their Solution*, American Mathematical Society, Providence, RI 1974. (Translated from the Russian.)

[5] D. J. SAKRISON, *Communication Theory: Transmission of Waveforms and Digital Information*, John Wiley, New York, 1968.

[6] H. TAUB AND D. L. SCHILLING, *Principles of Communication Systems*, McGraw-Hill, New York, 1971.

[7] R. P. FEINERMAN AND R. B. KELMAN, *The convergence of least squares approximations for dual orthogonal series*, Glasgow Math. J., 15 (1974), pp. 82–84; *Corrigenda*, Ibid., 184.

[8] N. I. ACHIEZER AND I. M. GLASMANN, *Theories der Linearen Operatoren im Hilbert-Raum*, Akademie-Verlag, Berlin, 1954. (In German, translated from the Russian.)

[9] E. A. CODDINGTON AND N. LEVINSON, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[10] R. P. FEINERMAN AND R. B. KELMAN, *Dual orthogonal series: An abstract approach*, Bull. Amer. Math. Soc., 81 (1975), pp. 733–736.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0003 $01.00/0

A TAUBERIAN REMAINDER THEOREM FOR THE HANKEL TRANSFORM*

WOLFRAM LUTHER†

Abstract. Using a general Parseval relation and the Wiener–Ganelius method, we give sharp Tauberian remainder results for the Hankel transform $F_\nu(x)=\int_0^\infty \sqrt{xu}J_\nu(xu)f(u)\,du$, $\nu \geqq -1/2$. The remainder of $f(u)$ covers the whole range between $o(1)$ and $O(u^{-1})$ which is a minorant for this transform. Applications to Fourier series and probability theory are possible.

1. Introduction. As an application of a general Parseval relation, Ridenhour and Soni [5] have recently obtained results of Tauberian character for the Hankel transform

$$F(x)=\int_0^\infty k(xu)f(u)\,du, \qquad k(u)=\sqrt{u}J_\nu(u)$$

under strong assumptions on $f(u)$. They derived their results from a Tauberian theorem of Karamata concerning the Laplace transform.

We will now give a more general Tauberian remainder theorem by generalizing a method introduced by T. Ganelius [3] to oscillating (symmetric) Fourier kernels $k(u)$ with some additional restrictions as in [6]. For simplicity however, we treat only the Hankel transform. Basic Tauberian theorems in [5] and [6] are therefore special cases of the remainder theorem given in this paper.

We cannot obtain precise remainder estimates with the method of Ridenhour and Soni, because the reciprocal Mellin transform $(k^M(t))^{-1}$ of the Laplace kernel increases exponentially, whereas the one of the Hankel kernel is of polynomial growth, and it is well known (see Lyttkens [4], Ganelius [2], Frennemo [1]) that the remainder of $f(u)$ depends on the growth of $(k^M(t))^{-1}$ in a neighborhood of the imaginary axis. Moreover, for the Laplace transform there is a critical order of decrease of the remainder, but for the Hankel transform and also for symmetric Fourier kernels the remainder of $f(u)$ has a minorant $M(u)=M_1u^{-1}$. The crucial step in the generalized method of Ganelius, an inversion of integration [3, p. 19 in the middle], is equivalent to the Parseval relation.

2. Basic assumptions. We use the assumptions on $f(u)$ from [5]:

(1)
- a) $f(u)\in \mathscr{P}_{\nu+1/2}$, i.e. $u^{\nu+1/2}f(u)\in L(0,R)$ for each finite $R>0$.
- b) $f(u)\in BV[a,\infty)$ for some $a>0$.
- c) $f(u)\to 0$ as $u\to\infty$.
- d) $|f(u)|\leqq c_1u^{\alpha_1}$, $\alpha_1\in \mathbf{R}$, $0<u\leqq a$.

Part b) can be replaced by $f(u)$ bounded in $[a,\infty)$, $f(u)\in L^2[a,\infty)$ and the assumption that $\int_0^\infty k(xu)f(u)\,du$ converges for $x>0$.

3. Main results. Now let us turn to the general remainder theorem:

THEOREM. *Assume that $f(u)$ satisfies the conditions* (1). *Furthermore,*

$$F_\nu(x)=\int_0^\infty k_\nu(xu)f(u)\,du=O\left(x^\gamma L(1/x)\exp\left(-W\left(-\frac{\log x}{2\pi}\right)\right)\right) \quad \text{as } x\to 0+,$$

* Received by the editors September 20, 1976.

† Lehrstuhl I für Mathematik, Rheinisch-Westfälische Technische Hochschule Aachen, Aachen, West Germany.

$-1<\gamma<\nu+1/2$, $\nu \geqq -1/2$, $k_\nu(u)=\sqrt{u}J_\nu(u)$, *with* $L(x)$ *slowly varying in the sense of Karamata* [5] *and* $W(x)$ *positive, increasing, subadditive, such that* $\lim_{x\to\infty} W(x)/x = 2\pi b$.

If $f(u)$ *satisfies the Tauberian condition*

$$\sup_{u_0 \leqq u \leqq v \leqq u(1+R(u))} (vf(v)-uf(u)) = O(u^{-\gamma}L(u)R(u))$$

with

$$R(u)=\exp\left(-\frac{W((\log u)/(2\pi))}{b+\gamma+2+\varepsilon}\right) \quad \textit{for an } \varepsilon>0,$$

then

$$f(u)=O(u^{-\gamma-1}L(u)R(u)) \quad \textit{as } u\to\infty.$$

Remark 1. The "O"-result can be replaced by the corresponding "o"-result. To derive the $o(1)$-Tauberian theorem of Ridenhour and Soni [5, Thm. 7], the assumption d) is not necessary, since a modification of $f(u)$ in $[0, R]$ does not change the main term. (For the Tauberian condition cf. Remark 4.)

Remark 2. Applying the Banach space-method of Ganelius [3, p. 43, Example 3, p. 45] to $k_\nu(\exp(-2\pi x))$, we can show that for $b+\gamma<\nu+1/2$ the remainder term cannot be substantially improved. *A minorant for* $R(u)$ *is* $M(u)=M_1u^{-1}$, *as is demonstrated by the example* $f(u)=J_\nu(u)J_\mu(u/2)u^{1/2-\mu}$ $(\mu \geqq 1/2)$ *with* $F_\nu(x)=0$ *for* $x<1/2$.

Proof. We generalize the method of Ganelius [3, pp. 18–20, pp. 34–40]. First we transform the integral to convolution form.

$$\begin{aligned}\psi_\nu(x) &:= F_\nu(\exp(-2\pi x))\\ &= 2\pi\int_{-\infty}^{\infty} k_\nu(\exp(-2\pi(x-y)))f(\exp(2\pi y))\exp(2\pi y)\,dy = K_\nu \times \phi(x)\end{aligned}$$

with $\phi(x)=\exp(2\pi x)f(\exp(2\pi x))$ and $K_\nu(x)=2\pi k_\nu(\exp(-2\pi x))$.

$$\hat{K}_\nu(t) := \int_{-\infty}^{\infty} \exp(-2\pi i x t)K_\nu(x)\,dx = 2^{it-1/2}\cdot\frac{\Gamma(1/4+\nu/2+it/2)}{\Gamma(3/4+\nu/2-it/2)}$$

for $-1<\operatorname{Im} t<\nu+1/2$.

$g(t) := (\hat{K}_\nu(t))^{-1}$ is holomorphic in $-3/2-\nu<\operatorname{Im} t<\infty$, and there holds $g(t)=O((1+|t|)^{\operatorname{Im} t+1/2})$. Furthermore,

$$\hat{K}_\nu(t)\cdot\hat{K}_\nu(-i-t)=1 \quad \text{in } -1<\operatorname{Im} t<\nu+1/2.$$

We put [3, p. 19]

$$Q^F(t) := g(t)\int_{-\infty}^{\infty} \hat{E}(t-y)\hat{\chi}(y)\exp(-2\pi i\eta y)\,dy$$

with

$$E(y)=\exp(-\pi y^2), \quad \chi(y)=\Omega(\pi y\Omega)^{-2}(\sin(\pi y\Omega))^2.$$

To obtain $\int_{-\infty}^{\infty}\exp(2\pi i t x)Q^F(t)\,dt$ we consider

$$\begin{aligned}Q^F_\alpha(t) &= g(t-i\alpha)\int_{-\infty}^{\infty} \hat{E}(t-y)\hat{\chi}(y)\exp(-2\pi i\eta y)\,dy\\ &= \hat{K}_\nu(-i(1-\alpha)-t)\cdot\hat{H}(t),\end{aligned}$$

with $H(y)=E(y)\chi(y-\eta)$ $(\in L(\mathbf{R}))$. For $0<\alpha<1$ holds

$$\hat{K}_\nu(-i(1-\alpha)-t)=\int_{-\infty}^{\infty} \exp(-2\pi ix(t+i(1-\alpha)))K_\nu(-x)\,dx$$

$$=\int_0^\infty u^{-\alpha-it}\cdot k_\nu(u)\,du.$$

We have $\int_\lambda^\mu u^{-\alpha-it}\cdot k_\nu(u)\,du=O(1+|t|)$ *for all* λ and μ.

Now we can apply Theorem 39 of [7] (concerning the Fourier transform of a convolution) to

$$\hat{K}_\nu(-i(1-\alpha)-t) \quad \text{and} \quad \hat{H}_x(t)=\exp(2\pi ixt)\hat{H}(t)=H(x+y)\hat{}\,.$$

With $I_\alpha(x)=2\pi\exp(2\pi(\tfrac{3}{2}-\alpha)x)J_\nu(\exp(2\pi x))$ it follows that

$$\int_{-\infty}^{\infty}\hat{K}_\nu(-i(1-\alpha)-t)\hat{H}_x(t)\,dt=\int_{-\infty}^{\infty}H(x-y)I_\alpha(y)\,dy.$$

For $\alpha\to 0$, an application of Lebesgue's convergence theorem yields

$$Q(x)=\int_{-\infty}^{\infty}H(x-y)I_0(y)\,dy,$$

and after a change of variables we have

$$xQ_1(x):=Q((\log x)/(2\pi))=x\int_0^\infty k_\nu(xu)H\Big(-\frac{1}{2\pi}\log u\Big)\,du.$$

$g(t)$ is holomorphic in the half plane $-\nu-3/2<\operatorname{Im} t<\infty$. Hence

$$Q(x)=\begin{cases}O(\exp(-2\pi\alpha_2 x)) & \text{as } x\to\infty,\ \alpha_2>0,\\ O(\exp(-2\pi(\nu+3/2-\alpha_3)|x|)) & \text{as } x\to-\infty,\ \alpha_3>0.\end{cases}$$

In a similar way we obtain $Q\times K_\nu=H$.

Finally we must legitimate the change of order of integration

$$Q\times\psi_\nu(x)=H\times\phi(x)\quad\text{for all } x.$$

By the Parseval relation [5, Thm. 1]:

$$\int_0^\infty F_\nu(x)Q_1(x)\,dx=\int_0^\infty f(t)H\Big(-\frac{1}{2\pi}\log t\Big)\,dt$$

and also

$$\int_0^\infty F_\nu(yu)Q_1(u)\,du=\int_0^\infty H\Big(-\frac{1}{2\pi}\log t\Big)f(t/y)(1/y)\,dt.$$

Substituting $t=\exp(-2\pi y)$, $y=\exp(-2\pi x)$, $u=\exp(2\pi y)$ we obtain

$$\int_{-\infty}^{\infty}H(y)f(\exp(2\pi(x-y)))\exp(2\pi(x-y))\,dy$$

$$=\phi\times H(x)=\int_{-\infty}^{\infty}F_\nu(\exp(-2\pi(x-y)))Q(y)\,dy=\psi_\nu\times Q(x).$$

From the assumptions in the theorem we find

$$|\psi_\nu(x)|\leqq c_3(\exp(-2\pi\gamma x)L(\exp(2\pi x))\exp(-W(x))),\qquad x\geqq x_0$$

and

$$\sup_{x_0/2 \leqq x \leqq y \leqq x + c_4 \exp(-W(x)/(b+\gamma+2+\varepsilon))} (\phi(y) - \phi(x)) = O\Big(\exp(-2\pi\gamma x) \cdot L(\exp(2\pi x)) \exp\Big(-\frac{W(x)}{b+\gamma+2+\varepsilon}\Big)\Big),$$

x_0 suitably chosen. Now it is obvious that the relation [3, p. 20 (1)]

$$|\phi(x)| \leqq 4 \cdot \sup_{0 \leqq v - y \leqq 2\Omega^{-1}} (\phi(x-y)E(y) - \phi(x-v)E(v)) + 6|Q \times \phi_\nu(x)|$$

still holds.

We estimate the first term on the right side by aid of the Tauberian condition analogous to [1, pp. 85–87]; for the second term we proceed as in [3, pp. 34–40] in connection with [6, Lemma 3]. This Lemma 3 characterizes the behavior of $L(x)$.

Beginning with the second term, we split $\psi_\nu(x)$ into

$$\psi_{\nu 1}(x) = \begin{cases} \psi_\nu(x), & x \geqq x_0, \\ 0, & x < x_0, \end{cases} \qquad \psi_{\nu 2}(x) = \begin{cases} 0, & x \geqq x_0, \\ \psi_\nu(x), & x < x_0. \end{cases}$$

By transforming the line of integration, we obtain for $x \geqq x_0$

$$|Q \times \psi_{\nu 1}(x)| = \Big|\Big(\int_{-\infty}^{0} + \int_{0}^{x-x_0}\Big)\psi_{\nu 1}(x-y)Q(y)\,dy\Big|$$
$$\leqq c_5 \Omega^{b+\gamma+1+\varepsilon} \exp(-2\pi\gamma x) L(\exp(2\pi x)) \exp(-W(x))$$

and

$$|Q \times \psi_{\nu 2}(x)| \leqq c_6 \Omega^{\alpha_2+1+\varepsilon} \exp(-2\pi\alpha_2 x), \qquad \alpha_2 \text{ large enough.}$$

In term 1 we write

$$\phi(x-y)E(y) - \phi(x-v)E(v) = (\phi(x-y) - \phi(x-v))E(y) + \phi(x-v)(E(y) - E(v)).$$

For $|v| \geqq x/2$ both terms on the right side are $O(\exp(-\pi x^2/8))$. Let us now assume $|v| < x/2$. $\exp(2\pi\lambda|v| - \pi v^2)$ is bounded for all real λ. Thus, by an interval-splitting and choosing

$$2(c_4\Omega)^{-1} = \exp(-W(x)/(b+\gamma+2+\varepsilon)),$$

we find

$$\sup_{0 \leqq v - y \leqq 2\Omega^{-1}} (\phi(x-y) - \phi(x-v))E(y) \leqq c_7 \exp(-2\pi\gamma x) L(\exp(2\pi x)) R(\exp(2\pi x)),$$
$$|v| < x/2, \quad x \geqq x_0$$

and then

$$|\phi(x)| \leqq c_8 \exp(-2\pi\gamma x) L(\exp(2\pi x)) R(\exp(2\pi x))$$
$$+ (8/\Omega) \sup_{|v| < x/2} |\phi(x-v)E'(\tilde{y})|, \qquad y < \tilde{y} < v, \quad x \geqq x_0.$$

Iteration gives the theorem.

Remark 3. *A special case is* $W(x) = 2\pi bx$. *Then*

$$F_\nu(x) = O(x^{b+\gamma} L(1/x)), \quad \text{as } x \to 0+,$$

implies

$$R(u) = u^{-b/(b+\gamma+2+\varepsilon)}.$$

In the case of more rapidly decreasing remainders with

$$W(x) = c_2 \exp(2\pi bx)$$

we obtain

$$R(u) = R_1 (\log u)^{1/b} u^{-1}.$$

The proof asks for modifying the considerations in [3, pp. 46–49]. For example put $\Omega = \xi = \zeta = \exp(2\pi x) \cdot (2\pi c_9 x)^{-1/b}$ for $2\pi y \geqq c_{10} + \log \zeta$ in the estimate of $Q(y)$, c_9, c_{10} sufficiently large.

Remark 4. *If we assume the modified Tauberian condition*

$$\sup_{u_0 \leqq u \leqq v \leqq u(1+R(u))} (f(v) - f(u)) = O(u^{-\gamma-1} L(u) R(u)),$$

it is more appropriate to use $E^k(y) = \exp(-\pi y^2 + 2\pi k y)$ *instead of* $E(y)$ *in the proof with* $k = 1$.

The method works also for more general Fourier kernels under the assumptions made in [6]. We need only that $g(t)$ has a holomorphic continuation with the same polynomial growth in a sufficiently large strip around the real axis.

In a further note we will study these questions and some applications.

REFERENCES

[1] L. FRENNEMO, *On general Tauberian remainder theorems*, Math. Scand., 17 (1965), pp. 77–88.

[2] T. GANELIUS, *The remainder in Wiener's Tauberian theorem*, Mathematica Gothoburgensia, 1 (1962), pp. 1–13.

[3] ———, *Tauberian Remainder Theorems*, Lecture Notes in Mathematics 232, Springer, Berlin–Heidelberg–New York, 1971.

[4] S. LYTTKENS, *The remainder in Tauberian theorems*, Ark. Mat., 2 (1954), pp. 575–588.

[5] J. R. RIDENHOUR AND R. P. SONI, *Parseval relation and Tauberian theorems for the Hankel transform*, this Journal, 5 (1974), pp. 809–821.

[6] K. SONI AND R. P. SONI, *Slowly varying functions and asymptotic behavior of a class of integral transforms, I, II, III*, J. Math. Anal. Appl., 49 (1975), pp. 166–179, 477–495, 612–628.

[7] E. C. TITCHMARSH, *Introduction to the Theory of Fourier Integrals*, Oxford University Press, London 1948.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0004 $01.00/0

ORTHOGONAL POLYNOMIALS THROUGH MOMENT GENERATING FUNCTIONALS*

ALLAN M. KRALL†

Abstract. It is shown that if the linear functional w generates moments $\{\mu_i\}_{i=0}^{\infty}$ through the formula $\mu_i = \langle w, x^i \rangle$, $i = 0, 1, \cdots$, then the Chebyshev polynomials $\{p_i(x)\}_{i=0}^{\infty}$ are orthogonal in the sense that $\langle w, p_m p_n \rangle = 0$ when $m \neq n$. In particular the Cauchy representations of the functionals associated with the Legendre, Jacobi and Bessel polynomials have this property when their action upon these polynomials is defined by a contour integral of sufficiently large radius.

Introduction. Quite recently there has been a flurry of activity in exploring orthogonality of polynomials in other than the classical sense. The author, with his student R. D. Morton [3], examined carefully a weight functional defined in terms of moments and derivatives of the Dirac delta function. Law and Sledd [5] used a recurrence relation to evaluate orthogonality and norms in a way independent of a weight function or measure. Al-Salam and Ismail [1] defined a discrete convolution orthogonality; and Jayne [2] used a complex Cauchy representation to achieve orthogonality, applying his results specifically to the Jacobi and Bessel polynomials.

The purpose of this note is to show that the classical Chebyshev polynomials $p_0 = 1$,

$$p_n = (1/\Delta_{n-1}) \begin{vmatrix} \mu_0 & \cdots & \mu_n \\ \vdots & & \vdots \\ \mu_{n-1} & \cdots & \mu_{2n-1} \\ 1 & \cdots & x^n \end{vmatrix}$$

where

$$\Delta_n = \begin{vmatrix} \mu_0 & \cdots & \mu_n \\ \vdots & & \vdots \\ \mu_n & \cdots & \mu_{2n} \end{vmatrix} \neq 0$$

are orthogonal with respect to any moment generating linear functional w defined on polynomials, i.e. any linear functional w satisfying

$$\langle w, x^i \rangle = \mu_i, \qquad i = 0, 1, \cdots.$$

Further, the norms squared of the Chebyshev polynomials are given by the usual formulas. In some instances these are analytic continuations of classical results. In other instances this kind of orthogonality is all that exists.

Orthogonality and norms.

THEOREM 1. *If $m \neq n$, then $\langle w, p_m p_n \rangle = 0$.*

The proof is virtually classical. Let $m < n$, and let $0 \leqq j \leqq m$. Then

$$\langle w, x^j p_n \rangle = (1/\Delta_{n-1}) \begin{vmatrix} \mu_0 & \cdots & \mu_n \\ \vdots & & \vdots \\ \mu_{n-1} & & \mu_{2n-1} \\ \langle w, x^j \rangle & \cdots & \langle w, x^{j+n} \rangle \end{vmatrix}$$

* Received by the editors October 12, 1976.

† Department of Mathematics, Pennsylvania State University, University Park, Pennsylvania, 16802.

$$= (1/\Delta_{n-1}) \begin{vmatrix} \mu_0 & \cdots & \mu_n \\ \vdots & & \vdots \\ \mu_{n-1} & & \mu_{2n-1} \\ \mu_j & \cdots & \mu_{n+j} \end{vmatrix}.$$

Since $0 \leqq j < n$, two rows in the determinant above are identical and 0 is the result. That p_m is orthogonal to p_n now follows by linearity.

THEOREM 2.

$$\langle w, p_0^2 \rangle = \mu_0, \qquad \langle w, p_n^2 \rangle = \Delta_n/\Delta_{n-1}, \qquad n = 1, 2, \cdots.$$

Proof. Note that

$$\langle w, p_n^2 \rangle = \langle w, x^n p_n \rangle + 0$$

because of orthogonality of p_n and powers of x of degree less than n. An observation of the formula in the proof of Theorem 1 shows

$$\langle w, x^n p_n \rangle = \Delta_n/\Delta_{n-1}.$$

Surprisingly the mere existence of w is sufficient to establish that a three term recurrence relation holds.

THEOREM 3. *If* $\langle w, x^i \rangle = \mu_i$, $i = 0, 1, \cdots$, *then there exist constants* B_n, C_n *such that*

$$p_{n+1}(x) = (x + B_n)p_n(x) - C_n p_{n-1}.$$

Again the proof is standard. Clearly

$$p_{n+1} = (x + B_n)p_n + \sum_{i=0}^{n-1} \alpha_i p_i.$$

If this is multiplied by p_m, $m < n-1$, and w is applied, the result is $\alpha_m \langle w, p_m^2 \rangle = 0$. Since $\langle w, p_m^2 \rangle \neq 0$, α_m is. Setting $\alpha_{n-1} = -C_n$ completes the formula.

By the same technique illustrated in [6] the constants B_n and C_n can be evaluated. If

$$p_n = x^n - S_n x^{n-1} + \cdots,$$

then

$$B_n = -S_{n+1} + S_{n-2},$$

and

$$C_n = \langle w, p_n^2 \rangle / \langle w, p_{n-1}^2 \rangle$$
$$= \Delta_n \Delta_{n-2} / \Delta_{n-1}^2.$$

The Cauchy representation. In terms of distribution theory, the Cauchy representation of a functional T is given by

$$\hat{T}(z) = \frac{1}{2\pi i} \left\langle T(x), \frac{1}{x-z} \right\rangle.$$

This formula implies in particular that for the Dirac delta function and its derivatives

$$\hat{\delta}(z) = (1/(2\pi i))/(-z),$$
$$\hat{\delta}^{(m)}(z) = (1/(2\pi i)) m!/(-z)^{m+1}.$$

Perhaps the most natural way to introduce the Cauchy representation in a way depending upon moments is to begin with the formula

$$w(x)=\sum_{n=0}^{\infty}(-1)^n\mu_n\,\delta^{(n)}(x)/n!,$$

discussed in [3], which in the classical cases is a restriction of the classical weight function. Then, formally, the Cauchy representation is

$$\hat{w}(z)=(1/(2\pi i))\sum_{n=0}^{\infty}(-1)^n\mu_n\,\hat{\delta}^{(n)}(z)/n!$$

$$=(1/(2\pi i))\sum_{n=0}^{\infty}-\mu_n/z^{n+1}.$$

The growth restriction originally imposed on the moments [3] was $|\mu_n|<cM^n n!$ This enabled the (inverse) Fourier transform to converge. In the present case, however, this is too lax, and, so, accordingly, in order to have convergence we now require that $|\mu_n|<cM^n$. $\hat{w}(z)$ then exists for $|z|>M$.

DEFINITION. Let $\langle\hat{w},\phi\rangle$ be given by the formula

$$\langle w,\phi\rangle=-\int\hat{w}(z)\phi(z)\,dz,$$

where the path of integration is a simple closed contour encircling the origin in a counterclockwise manner in the exterior of the circle $|z|=M$.

THEOREM 4. $\langle w,z^i\rangle=\mu_i$, $i=0,1,\cdots$. *Thus the Chebyshev polynomials* $\{p_i\}$ *are orthogonal with respect to* $\hat{w}$:

$$\langle\hat{w},p_np_m\rangle=-\int p_n(z)p_m(z)\hat{w}(z)\,dz=0$$

when $n\neq m$. *Further*

$$\langle w,p_0^2\rangle=\mu_0,$$

$$\langle w,p_n^2\rangle=\Delta_n/\Delta_{n-1}.$$

Examples. 1. *The Legendre polynomials.* The moments are $\mu_{2n}=1/(2n+1)$, $\mu_{2n+1}=0$, $n=0,\cdots$. Thus if $|z|>1$

$$\hat{w}(z)=-(1/(2\pi i))\sum_{m=0}^{\infty}(1/(2m+1)z^{2m+1})$$

$$=-(1/(2\pi i))\ln((z+1)/(z-1))^{1/2}$$

serves as a weight function.

2. *The Jacobi polynomials.* The moments for $\{P_n^{\alpha,\beta}(z)\}_{n=0}^{\infty}$ are given by

$$\mu_n=\sum_{j=0}^{n}\binom{n}{j}(-1)^j2^j(v)_j/(u+v)_j$$

$$=\sum_{j=0}^{n}\binom{n}{j}(-1)^{n-j}2^j(u)_j/(u+v)_j$$

where $\alpha=v-1$, $\beta=u-1$, and u, v and $u+v$ are not negative integers (see [3]). Jayne [2] has shown that the series for $\hat{w}(z)$ converges for $|z|>1$ when $u,v>0$. By using the regularization formula in [3], however, it can be shown that the series converges for

$|z|>1$ whenever u, v, $u+v$ are not negative integers. In fact,

$$\hat{w}(z)=(1+z)^{-1}{}_2F_1\left(\begin{matrix}u, 1\\ u+v\end{matrix}; \frac{2}{1+z}\right)=(z-1)^{-1}{}_2F_1\left(\begin{matrix}v, 1\\ u+v\end{matrix}; \frac{2}{1-z}\right).$$

3. *The Bessel polynomials.* The moments for the Bessel polynomials are $\mu_n = (-2)^{n+1}/(n+1)!$, $n=0, 1, \cdots$. Thus

$$\hat{w}(z)=(-1/(2\pi i))\sum_{n=0}^{\infty}((-2)^{n+1}/(n+1)!z^{n+1})$$
$$=(-1/(2\pi i))(e^{-2/z}-1)$$

outside any (arbitrarily small) circle centered at the origin. Since the last term (-1) may be ignored, orthogonality of the Bessel polynomials in the Cauchy sense is fully equivalent to the orthogonality introduced by H. L. Krall and O. Frink [4].

We remark that in the cases of the Laguerre and Hermite polynomials, the growth constraints placed on μ_n are not valid. And so, while the Cauchy representations for e^{-x} on $[0, \infty)$ and e^{-x^2} on $(-\infty, \infty)$ exist when $\operatorname{Re}(z)\neq 0$, the series representations are merely formal expansions. Either the Cauchy representations or the series expansions may be used to formally generate Cauchy orthogonality.

REFERENCES

[1] W. A. AL-SALAM AND M. E. H. ISMAIL, *Polynomials orthogonal with respect to discrete convolution*, J. Math. Anal. Appl., 55 (1976), pp. 125–139.
[2] J. W. JAYNE, *Polynomials orthogonal on a contour—The Bessel alternative*, manuscript.
[3] R. D. MORTON AND A. M. KRALL, *Distributional weight functions for orthogonal polynomials*, this Journal, 9 (1978), pp. 604–626.
[4] H. L. KRALL AND O. FRINK, *A new class of orthogonal polynomials: The Bessel polynomials*, Trans. Amer. Math. Soc., 65 (1949), pp. 100–115.
[5] A. G. LAW AND M. B. SLEDD, *Normalizing orthogonal polynomials by using their recurrence relation*, Proc. Amer. Math. Soc., 48 (1975), pp. 505–507.
[6] J. SHOBAT, *Notes on Chebyshev Polynomials, taken by I. M. Sheffer*, Dept. of Mathematics, Pennsylvania State Univ., University Park, 1960.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0005 $01.00/0

DISTRIBUTIONAL WEIGHT FUNCTIONS FOR ORTHOGONAL POLYNOMIALS*

ROBERT D. MORTON† AND ALLAN M. KRALL‡

Abstract. Given any collection of real numbers $\{\mu_i\}_{i=0}^{\infty}$, called moments, satisfying a Hamburger-like condition $\Delta_n = \det[\mu_{i+j}]_{i,j=0}^{n} \neq 0$ and a growth condition $|\mu_n| < cM^n n!$, where c, M are constant, $n = 0, 1, \cdots$, the Chebyshev polynomials $p_0 = 1$,

$$p_n(x) = [1/\Delta_{n-1}] \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \vdots & & & \vdots \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix},$$

$n = 1, 2, \cdots$, are shown to be orthogonal with respect to the linear functional

$$w(x) = \sum_{n=0}^{\infty} (-1)^n \mu_n \delta^{(n)}(x)/n!.$$

The problem of the existence of extensions of w to a space of test functions which includes polynomials is also discussed. It is shown that if $F^{-1}w(t)$ has an analytic continuation which has a classical Fourier transform, then that transform is the desired extension. If the continuation has an appropriate derivative which has a classical Fourier transform, then there exists a canonical regularizarion of a regular distribution which extends w.

As examples the Legendre, Jacobi, Laguerre, generalized Laguerre, Hermite and Bessel polynomials are offered. The Fourier transform establishes the connection between the functionals w and the classical weight functions when they exist. Further an extension of classical results is made in the cases of the generalized Laguerre and Jacobi polynomials. In the case of the Bessel polynomials, however, the measure of bounded variation, guaranteed by Boas's theorem, can only be found (?) as a Fourier transform, and so still remains an enigma.

I. DISTRIBUTIONAL WEIGHT FUNCTIONS

1. Introduction. Let $\psi(x)$ be a real analytic function whose Taylor's series converges to ψ for all x. Further let w be a linear functional acting on such functions which satisfies $\mu_n = \langle w, x^n \rangle$ for all $n = 0, 1, \cdots$. Then

$$\langle w, \psi \rangle = \langle w, \sum_{n=0}^{\infty} \psi^{(n)}(0) x^n/n! \rangle = \sum_{n=0}^{\infty} \psi^{(n)}(0) \langle w, x^n \rangle / n!.$$

Since $\psi^{(n)}(0)$ can be described by using the Dirac δ-function and its derivatives through

$$\psi^{(n)}(0) = (-1)^n \langle \delta^{(n)}, \psi \rangle,$$

it seems reasonable to expect

$$\langle w, \psi \rangle = \sum_{n=0}^{\infty} (-1)^n \langle \delta^{(n)}, \psi \rangle \cdot \langle w, x^n \rangle / n! = \left\langle \sum_{n=0}^{\infty} (-1)^n \mu_n \delta^{(n)}(x)/n!, \psi \right\rangle,$$

so that in the sense of distributions

$$w(x) = \sum_{n=0}^{\infty} (-1)^n \mu_n \delta^{(n)}(x)/n!,$$

which is a formula which has practical value provided the moments $\{\mu_i\}_{i=0}^{\infty}$ are known.

---

* Received by the editors February 19, 1976, and in final revised form September 1, 1976.

† Texas Instruments Corporation, Dallas, Texas 75222.

‡ Mathematics Department, Pennsylvania State University, University Park, Pennsylvania 16802.

Since for orthogonal polynomial sets, the moments are either given or can be calculated by techniques other than by using a weight function [4], the δ-expansion of w gives a computational method for calculating a "weight function" with respect to which these polynomials are indeed orthogonal.

When the moments $\{\mu_i\}_{i=0}^{\infty}$ are those associated with the various classical orthogonal polynomials, the Legendre polynomials, the Laguerre polynomials or the Hermite polynomials, the "weight functions" w yield virtually the same results concerning orthogonality and norms as the classical weight functions. Further, when the moments $\{\mu_i\}_{i=0}^{\infty}$ are those associated with the Jacobi polynomials, the generalized Laguerre polynomials or the Bessel polynomials, then w remains a suitable (distributional) "weight function" more or less regardless of the values of various parameters involved, even when a classical looking weight function cannot be found.

There are many additional questions concerning w which immediately present themselves. How far can this kind of weight function be extended (to how large a space of test functions)? When is w a *continuous* linear functional? What is the proper setting so that polynomials (on which it is obviously defined) are part of the space which is its domain? Then for various specific cases such as the Jacobi polynomials, the generalized Laguerre polynomials and the Bessel polynomials, what does the extension of w look like? How is it related to its classical counterpart when it exists?

The purpose of this article is to address these questions.

We make the *fundamental assumptions* that the moments $\{\mu_i\}_{i=0}^{\infty}$, are given, that

$$\Delta_n = \begin{vmatrix} \mu_0 & \cdots & \mu_n \\ \vdots & & \vdots \\ \mu_n & & \mu_{2n} \end{vmatrix} \neq 0$$

and that $|\mu_n| \leqq cM^n n!$, $n = 0, 1, \cdots$, for some arbitrary, but fixed, constants c and M.

Since it is crucial to what follows, we note that,

THEOREM 1.1. *The collections* $\{(-1)^m \delta^{(m)}(x)\}_{m=0}^{\infty}$ *and* $\{x^n/n!\}_{n=0}^{\infty}$ *form a biorthogonal set. That is,*

$$\langle (-1)^m \delta^{(m)}(x), x^n/n! \rangle = \begin{cases} 0 & \text{if } n \neq m, \\ 1 & \text{if } n = m. \end{cases}$$

We leave the proof to the reader.

2. The Spaces P and P'. The spaces D (infinitely differentiable functions with compact support), S (infinitely differentiable functions of rapid decay), E (infinitely differentiable functions with no growth restrictions) are well known, as are their dual spaces D' (no growth restrictions), S' (slow growth), E' (compact support) [2]. For our purposes, however, none of these pairs quite suit, since our function space should include polynomials, and, at the same time the dual space should include such functionals as those generated by exponential functions, i.e., without compact support. Therefore we introduce a new space P, which includes polynomials, satisfying

$$D \subset S \subset P \subset E.$$

The dual spaces then satisfy

$$D' \supset S' \supset P' \supset E'.$$

As we shall see, the connotation "slow growth" is appropriate for P, and "rapid decay" is appropriate for P' so that the analogies

D, E':	compact support,
D', E:	no restriction on growth,
P, S':	slow growth,
P', S:	rapid decay

are completed. Our immediate goal is to determine conditions under which the functional w can be continuously extended to P.

DEFINITION 2.1. We denote by P the linear vector space of all complex valued infinitely differentiable functions $\psi(x)$, $x \in E^1$, satisfying for all $\alpha > 0$ and $q \geqq 0$,

$$\lim_{|x|\to\infty} e^{-\alpha|x|}\psi^{(q)}(x) = 0.$$

We note that all polynomials with complex coefficients are in P.

DEFINITION 2.2. A sequence $\{\psi_j\}$ in P is said to *converge to zero in the sense of* P ($\psi_j \overset{P}{\to} 0$) provided for each $\alpha > 0$ and $q \geqq 0$ the sequence $\{e^{-\alpha|x|}\psi_j^{(q)}(x)\}$ converges to zero uniformly on, E^1. $\psi_j \overset{P}{\to} \psi_0$ if and only if $(\psi_j - \psi_0) \overset{P}{\to} 0$.

By way of comparing convergence in the spaces D, S, P, E we offer the following examples.

1. If
$$\psi_n(x) = \begin{cases} (1/n)\exp[-(1-x^2)^{-1}], & -1 < x < 1, \\ 0, & \text{otherwise}, \end{cases}$$
then $\psi_n \overset{D}{\to} 0$.

2. If $\psi_n(x) = (1/n)\,e^{-x^2}$, then $\psi_n \overset{S}{\to} O$, but $\{\psi_n\}$ does not converge in D.
3. If $\psi_n(x) = (1/n)x$, then $\psi_n \overset{P}{\to} 0$, but $\{\psi_n\}$ does not converge in D or S.
4. If $\psi_n(x) = (1/n)e^{x/n}$, then $\psi_n \overset{E}{\to} 0$, but $\{\psi_n\}$ does not converge in D, S or P.

DEFINITION 2.3. We denote by P' the space of continuous linear functionals on P.

We also use the word distribution to loosely describe an element in P', just as is done for elements in S' or E'.

3. A topology for *P*. Given a countable system of (semi)norms, $\|\cdot\|_1$, $\|\cdot\|_2, \cdots, \|\cdot\|_i, \cdots$, defined on a linear space Φ, a topology can be induced on to Φ by considering as open sets the collection

$$U_{p\varepsilon} = \{\psi : \psi \in \Phi, \|\psi\|_i < \varepsilon \text{ for } i \geqq p\}$$

and their translates, where $\varepsilon > 0$ and p is a nonnegative integer. These sets $U_{p\varepsilon}$ satisfy the properties of an open neighborhood basis of zero (see [2, p. 38]), and so Φ can be made into a linear topological space by taking these sets and their translates as a basis for the topology. Φ is said to be a *countably normed space.*

DEFINITION 3.1. Let $\psi \in P$. Then the pth norm of ψ is given by

$$\|\psi\|_p = \sup_x \{e^{-|x|/(a+1)}|\psi^{(q)}(x)|\}$$

for all a, q, integers, satisfying $0 \leqq a \leqq p$ and $0 \leqq q \leqq p$, $p = 0, 1, \cdots$.

THEOREM 3.2. *Let $\{\psi_j\}$ be a sequence of elements in P. Then $\psi_j \overset{P}{\to} 0$ if and only if $\psi_j \to 0$ in the sense of the countably normed topology induced by the norms of Definition 3.1.*

Proof. Assume $\psi_j \overset{P}{\to} 0$. Then, given $U_{p\varepsilon}$, we wish to show that there is a j_0 such that $j > j_0$ implies $\psi_j \in U_{p\varepsilon}$.

By assumption there exists a j_{aq} such that $j > j_{aq}$ implies

$$e^{-|x|/(a+1)}|\psi_j^{(q)}(x)| < \varepsilon.$$

Let $j_0 = \max_{a,q \leqq p} j_{aq}$. Then for $j > j_0$ we have

$$e^{-|x|/(a+1)}|\psi_j^{(q)}(x)| < \varepsilon$$

for all $a, q \leqq p$. That is, $\|\psi_j\|_p < \varepsilon$. Thus $\psi_j \in U_{p\varepsilon}$ for $j > j_0$. Thus $\psi_j \to 0$ in the countably normed topology.

Conversely assume $\psi_j \to 0$ in the topology. Then we wish to show that for each $a > 0$ and for each q there is a j_0 such that $j > j_0$ implies $e^{-|x|/(a+1)}|\psi_j^{(q)}(x)| < \varepsilon$.

Choose $\alpha = (a+1)^{-1}$ and $p = \max\{q, a\}$. By assumption there is a j_0 such that $j > j_0$ implies $\psi_j \in U_{p\varepsilon}$. Thus $\|\psi_j\|_p < \varepsilon$. This implies in particular that

$$\sup_x e^{-|x|/(a+1)}|\psi_j^{(q)}(x)| < \varepsilon,$$

and $\psi_j \overset{P}{\to} 0$.

We note that a continuous linear functional f on a countably normed space is continuous if and only if $\langle f, \psi_j\rangle \to 0$ whenever $\psi_j \to 0$ in the sense of the topology.

4. The spaces Z_M and Z. One of the major problems confronting us is the finding of a linear space upon which w is continuous. For example, if $\mu_n = n!$ (Laguerre moments) and $\psi(x) = e^{-x^2}$, which is certainly in P, then $\psi^{(2m)}(0) = (-1)^m(2m)!/m!$, and

$$\langle w, \psi\rangle = \sum_{m=0}^{\infty} (-1)^m(2m)!/m!,$$

which diverges.

Further, the action of w on a test function ψ intuitively implies

$$\langle w, \psi\rangle = \sum_{n=0}^{\infty} \mu_n \psi^{(n)}(0)/n! = \langle w, \sum_{n=0}^{\infty} \psi^{(n)}(0)x^n/n!\rangle,$$

suggesting that ψ should not only be infinitely differentiable, but analytic.

An obvious space to consider therefore is Z (see [3]), the space of Fourier transforms of elements in D. Surprisingly this is also slightly too large. Accordingly we turn our attention to a slightly smaller subspace $Z_{M\varepsilon}$.

DEFINITION 4.1. For $\varepsilon > 0$, let $Z_{M\varepsilon}$ be the subspace of all $\psi \in Z$ such that the support of $F^{-1}(\psi)$ is contained in the interval $[-(M+\varepsilon)^{-1}, (M+\varepsilon)^{-1}]$. (Equivalently let $Z_{M\varepsilon}$ be the space of all elements ψ such that

$$|x+iy|^q|\psi(x+iy)| < C_q e^{a|y|}$$

where $a \leqq M + \varepsilon$ (See [2, p. 971).)

LEMMA 4.2. *If $|\mu_n| \leqq cM^n \cdot n!$, $n = 0, 1, \cdots$, then $\sum_{n=0}^{\infty} |\mu_n||\psi^{(n)}(0)|/n!$ exists for all $\psi \in Z_{M\varepsilon}$.*

Proof. Let $\psi \in Z_M$ have inverse Fourier transform $\phi \in D$. Then

$$\psi(x) = \int_{-\infty}^{\infty} e^{ixt}\phi(t)\, dt,^{1}$$

which implies

$$|\psi^{(n)}(0)| \leqq \int_{-(M+\varepsilon)^{-1}}^{(M+\varepsilon)^{-1}} |t|^n |\phi(t)|\, dt \leqq (M+\varepsilon)^{-n} \int_{-(M+\varepsilon)^{-1}}^{(M+\varepsilon)^{-1}} |\phi(t)|\, dt.$$

Thus

$$\sum_{n=0}^{\infty} |\mu_n||\psi^{(n)}(0)|/n! \leqq c \sum_{n=0}^{\infty} \left(\frac{M}{M+\varepsilon}\right)^n \int_{-(M+\varepsilon)^{-1}}^{(M+\varepsilon)^{-1}} |\phi(t)|\, dt$$

$$\leqq c\left[1 - \frac{M}{M+\varepsilon}\right]^{-1} \int_{-(M+\varepsilon)^{-1}}^{(M+\varepsilon)^{-1}} |\phi(t)|\, dt < \infty.$$

THEOREM 4.3. *If* $|\mu_n| < cM^n \cdot n!$, $n = 0, 1, \cdots$, *then*

$$w(x) = \sum_{n=0}^{\infty} (-1)^n \mu_n \, \delta^{(n)}(x)/n!$$

is a continuous linear functional on $Z_{M\varepsilon}$ *in the sense of* Z.

Proof. According to Lemma 4.2 $\langle w, \psi\rangle$ is well defined for $\psi \in Z_{M\varepsilon}$. Suppose that $\psi_j \overset{Z}{\to} 0$. Then $\phi_j = F^{-1}\psi_j \overset{D}{\to} 0$. *Hence*

$$\langle w, \psi_j\rangle = \left| \sum_{n=0}^{\infty} \mu_n \psi^{(n)}(0)/n! \right| \leqq c\left[1 - \frac{M}{M+\varepsilon}\right]^{-1} \int_{-(M+\varepsilon)^{-1}}^{(M+\varepsilon)^{-1}} |\phi_j(t)|\, dt.$$

Since $\phi_j \overset{D}{\to} 0$, the integral approaches 0, and so does $\langle w, \psi_j\rangle$. Thus w is continuous.

5. Moments of extensions. We are faced with the problem of extending w from $Z_{M\varepsilon}$ to a larger space, such as P. For the moment, however, let us assume that a continuous extension of w, w_P, to P is possible. Since $x^n \notin Z_M$ for any $n = 0, 1, \cdots$, even though $\langle w, x^n\rangle = \mu_n$ is defined, it is not clear that $\langle w_P, x^n\rangle = \mu_n$. We show that this is indeed true.

THEOREM 5.1. *Let* w *have a continuous extension,* w_P, *acting on* P. *Then*

$$\langle w_P, x^n\rangle = \mu_n, \qquad n = 0, 1, \cdots.$$

Proof. (a) Let

$$\delta_m(t) = \begin{cases} A_m \exp\left[-\left(\dfrac{1}{m^2} - t^2\right)^{-1}\right], & \dfrac{-1}{m} < t < \dfrac{1}{m}, \\ 0, & \text{otherwise}, \end{cases}$$

where A_m is chosen so $\int_{-1/m}^{1/m} \delta_m(t)\, dt = 1$. These functions are infinitely differentiable. Further for $m \geqq M+\varepsilon$, the support of δ_m is within the interval $[-(M+\varepsilon)^{-1}, (M+\varepsilon)^{-1}]$. Finally it is evident that $\lim_{m\to\infty} \delta_m(t) = \delta(t)$.

---

[1] In order to conform with Bremermann [2] we use this form of the Fourier transform.

(b) Set

$$\psi_m(x) = F(\delta_m) = \int_{-1/m}^{1/m} e^{itx}\, \delta_m(t)\, dt.$$

By construction $\psi_m(x) \in Z_{M\varepsilon}$ when $m \geqq M + \varepsilon$.

Further, $\psi_m \overset{P}{\to} 1$. To see this, let N be the maximum of x in an arbitrary compact subset of E^1. Then, since

$$1 \geqq \cos tx \geqq 1 - x^2 t^2/2,$$

$$1 \geqq \int_{-1/m}^{1/m} [\cos tx]\, \delta_m(t)\, dt \geqq 1 - (x^2/2) \int_{-1/m}^{1/m} t^2\, \delta_m(t)\, dt$$
$$\geqq 1 - \frac{N^2}{2m^2} \int_{-1/m}^{1/m} \delta_m(t)\, dt = 1 - \frac{N^2}{2m^2}.$$

Further, since sin xt is odd,

$$\int_{-1/m}^{1/m} [\sin tx]\, \delta_m(t)\, dt = 0.$$

It follows, therefore, that

$$\int_{-1/m}^{1/m} e^{itx}\, \delta_m(t)\, dt$$

converges to 1 uniformly on the compact subset of E^1.

Similarly

$$\psi_m^{(k)}(x) = \int_{-1/m}^{1/m} (it)^k e^{itx}\, \delta_m(t)\, dt$$

converges to 0 uniformly on the compact set for all $k \geqq 1$. Since $\psi_m \in P$, it follows th $\psi_m \overset{P}{\to} 1$.

(c) Since w_P is assumed to be continuous on P,

$$\langle w_P, 1\rangle = \lim_{m\to\infty} \langle w_P, \psi_m\rangle = \lim_{m\to\infty} \langle w, \psi_m\rangle.$$

Note that

$$\langle w, \psi_m\rangle = \sum_{k=0}^{\infty} [(-1)^k \mu_k/k!] \left\langle \delta^{(k)}(x), \int_{-1/m}^{1/m} \delta_m(t)\, e^{itx}\, dt \right\rangle$$
$$= \sum_{k=0}^{\infty} [\mu_k/k!] \int_{-1/m}^{1/m} (it)^k\, \delta_m(t)\, dt.$$

And observe that if $m \geqq M + \varepsilon$

$$\left| \sum_{k=1}^{\infty} [\mu_k/k!] \int_{-1/m}^{1/m} (it)^k\, \delta_m(t)\, dt \right.$$
$$\leqq \sum_{k=1}^{\infty} \frac{\mu_k}{k!m^k} \int_{-1/m}^{1/m} \delta_m(t)\, dt = \sum_{k=1}^{\infty} \frac{|\mu_k|}{k!m^k}$$
$$\leqq \sum_{k=1}^{\infty} c(M/m)^k = (c/m)(M/[1 - M/m]).$$

Thus it follows that

$$\lim_{m\to\infty} \langle w, \psi_m\rangle = \mu_0.$$

(d) A multiplier for a space Φ is an infinitely differentiable function f such that if $\psi \in \Phi$, then $f\psi \in \Phi$, and if $\psi_j \overset{\Phi}{\to} 0$, then $f\psi_j \overset{\Phi}{\to} 0$. It is clear that x^n is a multiplier on P as well as on Z and Z_M.

In particular since $\psi_m \overset{P}{\to} 1$, then $x^n\psi_m \overset{P}{\to} x^n$. Hence

$$\langle w_P, x^n\rangle = \lim_{m\to\infty} \langle w_P, x^n\psi_m\rangle = \lim_{m\to\infty} \langle w, x^n\psi_m\rangle.$$

Therefore

$$\langle w, x^n\psi_m\rangle = \sum_{k=0}^{\infty} [(-1)^k\mu_k/k!]\left\langle \delta^{(k)}(x), x^n \int_{-1/m}^{1/m} \delta_m(t)\, e^{itx}\, dx\right\rangle$$

$$= \sum_{k=n}^{\infty} \left[\binom{k}{k-n} n!\mu_k/k!\right] \int_{-1/m}^{1/m} (it)^{k-n}\, \delta_m(t)\, dt.$$

Now note that

$$\left|\sum_{k=n+1}^{\infty} \frac{\mu_k}{(k-n)!} \int_{-1/m}^{1/m} (it)^{k-n}\, \delta_m(t)\, dt\right| \leqq \sum_{k=n+1}^{\infty} c\frac{M^k k!}{(k-n)!}(1/m)^{k-n}$$

$$= cM^n \frac{d^n}{dx^n}\left(\frac{x^{n+1}}{1-x}\right)\Bigg|_{x=M/m}.$$

(It is understood that $m > M+\varepsilon$ so $M/m < 1$.) In turn the nth derivative is equal to

$$[x/(1-x)][(n+1)! + P(x/(1-x))]|_{x=M/n}$$

where $P(y)$ is a polynomial in y of degree n with no constant term. The coefficients of P depend only on n. Hence it follows that

$$\sum_{k=n+1}^{\infty} c\frac{M^k k!}{(k-n)!}(1/m)^{k-n}$$

can be made arbitrarily small by choosing m sufficiently large. Therefore

$$\lim_{m\to\infty} \langle w, x^n\psi_m\rangle = \lim_{m\to\infty} \mu_n \int_{-1/m}^{1/m} \delta_m(t)\, dt = \mu_n,$$

and $\langle w_P, x^n\rangle = \mu_n$.

II. EXTENSIONS

6. The Fourier transform. It is our chief concern now to extend the functional w to act upon as large a space as possible, with the specific aim of extending w to act on P. Since our major tool in this extension process is the inverse Fourier transform, we formally introduce it at this point. We shall need an additional space of test functions in order to conveniently carry out our calculations.

DEFINITION 6.1. *For $\varepsilon > 0$, let $D_{M\varepsilon}$ be the subspace of all $\phi \in D$ such that the support of ϕ is contained in the interval $[-(M+\varepsilon)^{-1}, (M+\varepsilon)^{-1}]$.*

We note that the image of $D_{M\varepsilon}$ under the Fourier transform is $Z_{M\varepsilon}$. Likewise $F^{-1}(Z_{M\varepsilon}) = D_{M\varepsilon}$ [2, p. 97], just as is the case with D and Z: $F(D) = Z$, $F^{-1}(Z) = D$.

Since for $\phi \in D$ or $D_{M\varepsilon}$, $f \in D'$ or $D'_{M\varepsilon}$ the Fourier transform of f is defined through

$$\langle Ff, F\phi\rangle = 2\pi\langle f, \phi\rangle;$$

if $\psi(x) = \int_{\infty}^{\infty} e^{ixt}\phi(t)\,dt = F\phi$, and $Ff = g$, the inverse Fourier transform of $g \in Z'$ or $Z'_{M\varepsilon}$ will be given by

$$\langle F^{-1}g, F^{-1}\psi\rangle = \frac{1}{2\pi}\langle g, \psi\rangle,$$

where $\psi \in Z$ or $Z_{M\varepsilon}$, and

$$\phi(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ixt}\psi(x)\,dx = F^{-1}\psi$$

is in D or $D_{M\varepsilon}$.

It is apparent that F^{-1} is a bijective mapping of Z' or $Z'_{M\varepsilon}$ onto D' or $D'_{M\varepsilon}$ under which the usual formulas hold:

1. $(F^{-1}g)^{(n)} = F^{-1}((-ix)^n g)$,
2. $F^{-1}(g^{(n)}) = (it)^n F^{-1}g$,

and in particular,

3. $F^{-1}(\delta^{(n)}) = (it)^n/(2\pi)$.

7. The extensions to *Z* and *S*. We observe that the inverse Fourier transform of w exists.

LEMMA 7.1.

$$F^{-1}w(t) = \frac{1}{2\pi}\sum_{n=0} \mu_n(-it)^n/n!.$$

Since by assumption $|\mu_n| < M^n n!$, $F^{-1}w$ represents an analytic function for $|t| < 1/M$. This is consistent with the requirement that $F^{-1}w$ be defined on $D_{M\varepsilon}$, i.e., on those distributions with support in the interval $[-(M+\varepsilon)^{-1}, (M+\varepsilon)^{-1}]$.

Note further $F^{-1}w$ is a regular functional, so that

$$\langle F^{-1}w, \phi\rangle = \int_{-\infty}^{\infty} \overline{F^{-1}w(t)}\phi(t)\,dt.$$

THEOREM 7.2. *w has extensions w_Z which are distributions on Z.*

Proof. Let g be any locally integrable extension of $f(t) = F^{-1}w(t)$, $t \in [-(M+\varepsilon)^{-1}, (M+\varepsilon)]$. Then $\langle g, \phi\rangle$ exists for all $\phi \in D$ since ϕ has compact support. We then define w_Z through the formula $w_Z = Fg$.

THEOREM 7.3. *w has extensions w_S which are distributions on S.*

Proof. Let g be any locally integrable extension of

$$f(t) = F^{-1}w(t), \qquad t \in [-(M+\varepsilon)^{-1}, (M+\varepsilon)^{-1}],$$

which grows no faster than a polynomial as $|t| \to \infty$. Then $w_Z = Fg$ is a regular distribution on S and clearly extends w.

8. The extensions to *P* and *E*. We are faced with an abundance of extensions from $Z_{M\varepsilon}$ to Z and from $Z_{M\varepsilon}$ to S. The question remaining is which, if any, can be further extended to P or to E? Certainly there is no reason—a priori—for such an extension to exist.

We can gain some insight into what occurs if we first enlarge M to $2M$ and then examine the procedure of extending w from $Z_{2M\varepsilon}$ to $Z_{M\varepsilon}$. On $Z_{2M\varepsilon}$ w has an inverse Fourier transform which is analytic when $|t|<(2M)^{-1}$, while on $Z_{M\varepsilon}$, w has an inverse Fourier transform which is analytic when $|t|<M^{-1}$. Clearly the latter is an analytic continuation of the former. An application of the Fourier transform gives the desired extension.

In order to use analytic continuation to extend w further, however, additional assumptions will be required.

LEMMA 8.1. *Let $f(z)$ be analytic in the region $|\mathrm{Im}(z)|<s_0$ with $|f(z)|\leqq h_0(t)$, $|f'(z)|\leqq h_1(t)$ for all $z=t+is$ with $|s|<s_0$. Further, assume that $\lim_{|t|\to\infty} h_0(t)=0$, and that $\int_{-\infty}^{\infty} h_1(t)\,dt<\infty$. Then $f(t)$ has a classical Fourier transform $g(x)$. Further, there exist constants G and $r>0$ such that*

$$|g(x)|<Ge^{-r|x|}/|x|.$$

Proof. (a) Choose r, ε from the open interval $(0, s_0)$, and consider the contour C shown in Fig. 1.

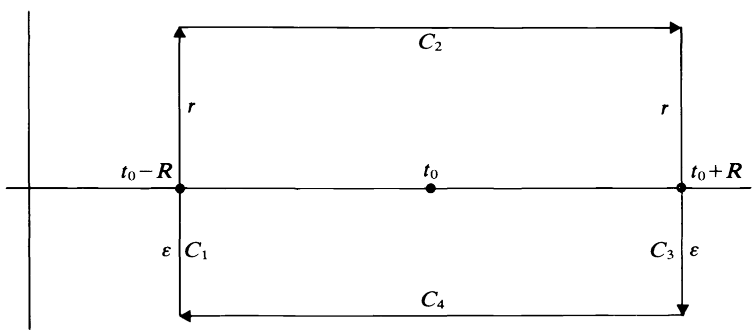

FIG. 1

Let $x>0$. Since $f(z)\,e^{ixz}$ is analytic in and on C, we apply Cauchy's integral theorem,

$$\int_C f(z)\,e^{ixz}\,dz=\sum_{k=1}^{4}\int_{C_k} f(z)\,e^{ixz}\,dz=0.$$

On C_1 the real part of z is fixed at t_0-R. Hence $|f(z)|<h(t_0-R)$ and

$$\left|\int_{C_1} f(z)\,e^{ixz}\,dz\right|\leqq\int_{-\varepsilon}^{r}|f(z)|\,e^{-sx}\,ds$$

$$\leqq h_0(t_0-R)\int_{-\varepsilon}^{r} e^{-sx}\,ds=h_0(t_0-R)[(e^{\varepsilon x}-e^{-rx})/x].$$

As $R\to\infty$, this approaches 0, since $\lim_{R\to\infty} h_0(t_0-R)=0$.

On C_3 we also find

$$\left|\int_{C_3} f(z)\,e^{ixz}\,dz\right|\leqq\int_{-\varepsilon}^{r}|f(z)|\,e^{-sx}\,ds\leqq h_0(t_0+R)[(e^{\varepsilon x}-e^{-rx})/x],$$

which approaches 0, since $\lim_{R\to\infty} h_0(t_0+R)=0$.

Therefore as $R\to\infty$, we find

$$\int_{-\infty}^{\infty} f(t+ir)\,e^{ix(t+ir)}\,dt=\int_{-\infty}^{\infty} f(t-i\varepsilon)\,e^{ix(t-i\varepsilon)}\,dt.$$

(b) Let $\varepsilon = 0$, $r > 0$. Integrating by parts, we find

$$\int_{-\infty}^{\infty} f(t+ir)\, e^{ixt}\, dt = (1/ix)\Big[f(t+ir)\, e^{ixt}\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f'(t+ir)\, e^{ixt}\, dt\Big].$$

Since $|f(t+ir)| \leqq h_0(t)$ *and* $\lim_{|t|\to\infty} h_0(t) = 0$,

$$\Big|\int_{-\infty}^{\infty} f(t+ir)\, e^{ixt}\, dt\Big| = (1/|x|)\Big|\int_{-\infty}^{\infty} f'(t+ir)\, e^{ixt}\, dt\Big|,$$

$$\leqq (1/|x|)\int_{-\infty}^{\infty} |f'(t+ir)|\, dt$$

$$\leqq (1/|x|)\int_{-\infty}^{\infty} h_1(t)\, dt.$$

It follows, therefore, that

$$\Big|\int_{-\infty}^{\infty} f(t)\, e^{ixt}\, dt\Big| \leqq Ge^{-rx}/|x|.$$

(c) Let $x < 0$. By interchanging ε and r and then setting $\varepsilon = 0$, $r > 0$, we show by a similar argument that

$$\Big|\int_{-\infty}^{\infty} f(t)\, e^{ixt}\, dt\Big| \leqq Ge^{rx}/|x|.$$

(d) These estimates when combined show that $f(t)$ has a classical Fourier transform $g(x)$ and that $|g(x)| < Ge^{-r|x|}/|x|$.

We can now consider the possibility of extending w to P.

THEOREM 8.2. *Let $f(z)$ be the analytic continuation of $F^{-1}w$, where $w = \sum_{n=0}^{\infty} (-1)^n \mu_n \delta^{(n)}(x)/n!$ is a weight distribution on Z_{M_ε}. Assume that for $z = s + it$, $f(z)$ satisfies the following*:

1. *$f(z)$ is analytic in the strip $|\mathrm{Im}\,(z)| = |s| < s_0$ for some $s_0 > 0$.*

2. *When $|s| < s_0$, $|f(z)| \leqq h_0(t)$ and $|f'(z)| \leqq h_1(t)$, where* $\lim_{|t|\to\infty} h_0(t) = 0$, $\int_{-\infty}^{\infty} h_1(t)\, dt < \infty$. *Then the classical Fourier transform of $f(t)$, $w_P(x)$, is a continuous linear functional on P and is an extension of w.*

Proof. By Lemma 8.1 $f(t)$ has a Fourier transform $g(x) = \int_{-\infty}^{\infty} f(t)\, e^{ixt} dt$ which satisfies, for all x, $|g(x)| < Ge^{-r|x|}/|x|$ for some constants G and $r > 0$. Since for $\psi \in P$

$$|\langle g, \psi\rangle| = \Big|\int_{-\infty}^{\infty} \overline{g(x)}\psi(x)\, dx\Big|$$

$$< G \sup_{x\in(-\infty,\infty)} [e^{-(r/2)|x|}|\psi(x)|] \int_{-\infty}^{\infty} e^{-(r/2)|x|}\, dx < \infty,$$

it follows that $\langle g, \psi\rangle$ not only exists but is a continuous linear functional. Thus $w_P(x) = g(x)$ is the desired extension.

COROLLARY 8.3. *Let the extension, w_P, of Theorem 8.2. have compact support. Then w_P can be extended to a unique continuous linear functional, w_E, on E.*

Proof. Since $w_P \in P' \subset D'$, according to Bremermann [2, p. 27], w_P possesses a unique extension in E'.

III. REGULARIZATIONS

9. An example. We have just seen how w can be extended when the analytic continuation, f, of $F^{-1}w$ has a classical Fourier transform. In what follows we shall devote our attention to the problem of extending w when a classical Fourier transform does not exist.

As an illustration of what can occur consider the case of the generalized Laguerre polynomials $\{L_n^{-3/2}\}_{n=0}^{\infty}$. In this case

$$w(x)=\sum_{n=0}^{\infty}(-1)^n\frac{\Gamma(n-\frac{1}{2})}{\Gamma(-\frac{1}{2})n!}\delta^{(n)}(x)$$

and

$$F^{-1}w(t)=\frac{1}{2\pi}\sum_{n=0}^{\infty}\frac{\Gamma(n-\frac{1}{2})}{\Gamma(-\frac{1}{2})n!}(-it)^n.$$

It is easy to see that the analytic continuation of $F^{-1}w$ is $(1/(2\pi))(1+it)^{1/2}$, which does not have a classical Fourier transform.

It does, however, have a derivative, $f'(t)=[i/(4\pi)](1+it)^{-1/2}$, which is classically transformable:

$$g(x)=F[f'(t)]=\begin{cases}\dfrac{ix^{-1/2}e^{-x}}{\Gamma(-\frac{1}{2})}, & x>0,\\ 0, & x\leqq 0.\end{cases}$$

Further, since

$$\langle F[f'], F[\phi]\rangle=\langle -ixF[f], F[\phi]\rangle$$

with the extension w_P equal to $F[f]$, we see that w_P and $g(x)/(ix)$ should represent the same linear functional. But

$$w_P(x)=g(x)/ix=\begin{cases}\dfrac{x^{-1/2}e^{-x}}{\Gamma(-\frac{1}{2})}, & x>0,\\ 0, & x\leqq 0\end{cases}$$

does not generate a continuous linear functional even on Z due to the singularity at $x=0$.

The difficulty exhibited above can be circumvented by a process known as *regularization* (see [3, pp. 45–81]), and it is this procedure which (still) generates continuous extensions of w to P or E.

A close examination of regularization follows. Let us say in closing this section that the regularization of our example leads to the functional w_P generated by

$$\langle w_P, \psi\rangle=(1/\Gamma(-\tfrac{1}{2}))\int_0^{\infty}x^{-3/2}(e^{-x}\psi(x)-\psi(0))\,dx,$$

or $w_P=e^{-x}x_+^{-3/2}/\Gamma(-\frac{1}{2})$, where

$$\langle x_+^{-3/2}, \psi\rangle=\int_0^{\infty}x^{-3/2}(\psi(x)-\psi(0))\,dx.$$

10. Regularizations. For the moment let us consider the regularization of $h(x)$ which has a singularity only at 0. We assume that $h(x)$ is integrable over every bounded region of E' not containing 0 either in its interior or on its boundary. A regularization of $h(x)$ is a continuous linear functional which coincides with $h(x)$ except at 0. That is, for every ψ in the test space which vanishes in a neighborhood of 0, the functional has the same value as $\int_{-\infty}^{\infty}\overline{h(x)}\psi(x)\,dx$.

We quote the following two theorems from [3, p. 11].

THEOREM 10.1. *If there exists an* $m>0$ *such that* $x^m h(x)$ *is locally integrable, then* $h(x)$ *can be regularized.*

THEOREM 10.2. *Any two regularizations of* $h(x)$ *differ by a functional concentrated at* 0.

If the regularization of $h(x)$ also preserves the operations of addition, multiplication by an appropriate function and differentiation, then the regularization is called *canonical.* Following [3] we write $h=\mathrm{CR}h(x)$ to denote that h is the canonical regularization of $h(x)$.

Finally we shall restrict our attention to those functions which can be written as

$$h(x)=\sum p_i(x)q_i(x)$$

where each p_i is infinitely differentiable, and each q_i is one of the functions x_+^λ, x_-^λ and x^{-n}. Then

$$\mathrm{CR}h(x)=\sum p_i(x)\mathrm{CR}q_i(x).$$

We remark that if $h(x)$ has singularities of the kind mentioned above at more than one point, say at $x_0<x_1\cdots<x_n$, and if $y_1,\cdots,y_n$ are chosen so $x_0<y_1<x_1\cdots<y_n<x_n$, $y_0=-\infty$, $y_{n+1}=\infty$, and

$$h_i(x)=\begin{cases}h(x), & y_{i-1}<x<y_i,\\ 0, & \text{otherwise},\end{cases}$$

then $h(x)$ can be decomposed into the sum $\sum h_i(x)$, and each term can be handled separately. This situation arises in a discussion of Jacobi polynomials, where singularities occur at ± 1. For our purposes here we shall restrict our attention to singularities at 0 only.

11. Regularized extensions. We assume that $f(z)$ is the analytic continuation of $F^{-1}w$ and that $f^{(m)}(z)$, $z=t+is$, is analytic with $|f^{(m)}(z)|\leqq h_0(t)$, $|f^{(m+1)}(z)|\leqq h_1(t)$ when $|s|<s_0$, where $\lim_{|t|\to\infty}h_0(t)=0$ and $\int_{-\infty}^{\infty}h_1(t)\,dt<\infty$.

THEOREM 11.1. *Let* $g(x)$ *denote the classical Fourier transform of* $f^{(m)}(t)$. *Assume that* $g(x)/(-ix)^m$ *has a canonical regularization* $h(x)$. *Then there exist constants* c_k, $k=0,\cdots,m-1$, *such that* $w_P(x)=h(x)+\sum_{k=0}^{m-1}c_k\,\delta^{(k)}(x)$ *is a continuous linear extension of* w *to* P.

Proof. The constraints on f guarantee that $\int_{-\infty}^{\infty}g(x)\psi(x)\,dx$ is continuous on P. This, in turn, insures that the canonical regularization of $g(x)/(-ix)^m$ will be continuous and linear on P.

Now $w_P=Ff$ is an extension of w. We claim there exist constants c_k, $k=0,\cdots,m-1$, such that

$$w_P(x)=h(x)+\sum_{k=0}^{m-1}c_k\,\delta^{(k)}(x).$$

Since h is continuous and linear on P, so will w_P be. Thus w_P is an extension of w to P.

To see that in fact

$$w_P(x)=Ff(x)=h(x)+\sum_{k=0}^{m-1}c_k\,\delta^{(k)}(x),$$

we proceed as follows.

(a) Let $\gamma(t)=F^{-1}h(t)$. Then

$$2\pi\langle\gamma^{(m)},\phi(t)\rangle=\langle F\gamma^{(m)},F\phi\rangle=\langle(-ix)^m h,F\phi\rangle.$$

Since $(-ix)^m$ is infinitely differentiable,

$$\begin{aligned}(-ix)^m h(x)&=(-ix)^m \mathrm{CR}[g(x)/(-ix)^m]\\&=\mathrm{CR}[(-ix)^m g(x)/(-ix)^m]=g(x).\end{aligned}$$

Thus

$$2\pi\langle\gamma^{(m)},\phi\rangle=\langle g,F\phi\rangle=2\pi\langle f^{(m)},\phi\rangle.$$

That is, $\gamma^{(m)}=f^{(m)}$.

(b) Since every distribution in D' has antiderivatives of mth order, we conclude

$$f(t)=\gamma(t)+\sum_{k=0}^{m-1}\frac{c_k}{2\pi}(it)^k.$$

Taking Fourier transforms, we find

$$w_P(x)=h(x)+\sum_{k=0}^{m-1}c_k\,\delta^{(k)}(x).$$

COROLLARY 11.2. *The coefficients* c_k, $k=0,1,\cdots,m-1$, *are given by*

$$c_k=\frac{(-1)^k}{k!}[\mu_k-\langle h,x^k\rangle].$$

Proof. Since $\langle w_P,x^k\rangle=\mu_k$, $k=0,\cdots$, and $\langle\delta^{(l)}(x),x^k\rangle=(-1)^k k!\,\delta_{kl}$,

$$\begin{aligned}\mu_k=\langle w_P,x^k\rangle&=\langle h,x^k\rangle+\sum_{k=0}^{m-1}c_l\langle\delta^{(l)}(x),x^k\rangle\\&=\langle h,x^k\rangle+(-1)^k k!c_k.\end{aligned}$$

COROLLARY 11.3. *Let the extension* w_P *of Theorem* 11.1 *have compact support. Then* w_P *can be extended to a unique continuous linear functional,* w_E, *on* E.

IV. ORTHOGONAL POLYNOMIALS

12. General orthogonal polynomials. Let us consider the polynomials $p_n(x)$ defined by $p_0=1$,

$$p_n(x)=\frac{1}{\Delta_{n-1}}\begin{vmatrix}\mu_0&\mu_1&\cdots&\mu_n\\ \mu_1&\mu_2&\cdots&\\ &&&\\ \mu_{n-1}&\mu_n&\cdots&\mu_{2n-1}\\ 1&x&\cdots&x^n\end{vmatrix},$$

$n=1,2,\cdots$, where

$$\Delta_n=\begin{vmatrix}\mu_0&\cdots&\mu_n\\ \vdots&&\\ \mu_n&\cdots&\mu_{2n}\end{vmatrix}\neq 0,\qquad n=0,1,\cdots.$$

We assume that the distribution w has been extended to w_P, which is continuous on P.

THEOREM 12.1. *The collection* $\{p_n(x)\}_{n=0}^{\infty}$ *is mutually orthogonal with respect to* w_P. *That is, if* $m \neq n$, $\langle w_P, p_m p_n\rangle = 0$.

Proof. We note that if $k < n$, then

$$\langle w_P, x^k p_n(x)\rangle = \frac{1}{\Delta_{n-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \mu_{n-1} & \mu_n & & \mu_{2n-1} \\ \mu_k & \mu_{k+1} & \cdots & \mu_{k+n} \end{vmatrix} = 0,$$

since the last row will be identical with one above. Thus, when $m < n$, if $p_m(x) = \sum_{k=0}^{m} c_k x^k$, then

$$\langle w_P, p_m p_n\rangle = \sum_{k=0}^{m} c_k \langle w_P, x^k p_n\rangle = 0.$$

THEOREM 12.2. $\langle w_P, p_n^2\rangle = \Delta_n/\Delta_{n-1} \neq 0$.

Proof. $\langle w_P, p_n^2\rangle = \langle w_P, x^n p_n\rangle$, which, by observation of the formula above, is Δ_n/Δ_{n-1}.

The precise connection between w_P and the classical weight functions follows.

13. The classical orthogonal polynomials.

A. The Legendre polynomials. The moments for the Legendre polynomials are $\mu_{2n} = 2/(2n+1)$, $\mu_{2n+1} = 0$. Thus

$$w = \sum_{n=0}^{\infty} \frac{2\delta^{(2n)}(x)}{(2n+1)!},$$

and

$$F^{-1}w = -\frac{1}{\pi} \sum_{n=0}^{\infty} \frac{(it)^{2n}}{(2n+1)!}.$$

This is a power series representation for

$$f(t) = \frac{(e^{it} - e^{-it})}{(2\pi it)} = \frac{[\sin t]}{(\pi t)}.$$

This function has a classical Fourier transform

$$w_E(x) = \begin{cases} 1, & -1 \leqq x \leqq 1, \\ 0, & |x| > 1. \end{cases}$$

B. The Laguerre polynomials. The moments for the Laguerre polynomials are $\mu_n = n!$. Thus

$$w = \sum_{n=0}^{\infty} (-1)^n \delta^{(n)}(x),$$

and

$$F^{-1}w = \sum_{n=0}^{\infty} (-it)^n/(2\pi).$$

This power series representation converges when $|t|<1$ to $(1/2\pi)(1+it)^{-1}$. Thus the analytic continuation of $F^{-1}w$ is

$$f(z)=\frac{1}{2\pi}(1+iz)^{-1}.$$

We see with $s_0=\frac{1}{2}$, $h_0(t)=(1/(2\pi))(1/(4+t^2))^{-1/2}$, $h_1(t)=(1/(2\pi))(1/(4+t^2))^{-1}$, that $\lim_{|t|\to\infty} h_0(t)=0$ and that $\int_{-\infty}^{\infty} h_1(t)\,dt<\infty$. Thus $w_P=Ff$ extends w to P. Cauchy's residue then establishes that

$$w_P=\begin{cases} e^{-x}, & x>0,\\ 0, & x\leqq 0.\end{cases}$$

C. The Hermite polynomials. The moments for the Hermite polynomials are $\mu_{2n}=\sqrt{\pi}\,(2n)!/(4^n n!)$, $\mu_{2n+1}=0$. Thus

$$w=\sum_{n=0}^{\infty}\frac{\sqrt{\pi}\,\delta^{(2n)}(x)}{4^n n!}$$

and

$$F^{-1}w=(2\sqrt{\pi})^{-1}\sum_{n=0}^{\infty}\frac{(-t^2/4)^n}{n!}.$$

This is the power series representation for

$$f(z)=(2\sqrt{\pi})^{-1}\,e^{-z^2/4}.$$

With $s_0=1$, $h_0(t)=e^{-t^2/4}$, $h_1(t)=[|t|/2]\,e^{-t^2/4}$, we satisfy the conditions to extend w, and

$$w_P=Ff(x)=e^{-x^2},\qquad -\infty<x<\infty.$$

14. Generalized Laguerre polynomials. H. L. Krall [4] has shown that the differential equation

$$(l_{22}x^2+l_{21}x+l_{20})p_n+(l_{11}x+l_{10})p_n=(l_{11}n+l_{22}n(n-1))p_n$$

has a polynomial solution $p_n(x)$ of degree n for each $n=0,1,\cdots$, if and only if the moments $\{\mu_i\}_{i=0}^{\infty}$ satisfy $\Delta_n=0$ and

$$l_{11}\mu_n+l_{10}\mu_{n-1}+(n-1)[l_{22}\mu_n+l_{21}\mu_{n-1}+l_{20}\mu_{n-2}]=0,$$

$n=1,2,\cdots$.

For the Laguerre equation

$$xL_n^{(\alpha)''}-[x-\alpha-1]L_n^{(\alpha)'}+nL_n^{(\alpha)}=0$$

the recurrence relation above is $\mu_n=[n+\alpha]\mu_{n-1}$. Consequently,

THEOREM 14.1. *Let* $\mu_0=1$. *Then when* $\alpha\neq -1,-2,\cdots$,

$$\mu_n=\frac{\Gamma(n+\alpha+1)}{\Gamma(\alpha+1)}.$$

Note that these moments have been calculated by a technique *not* dependent on the existence of a weight function. Further when α is a negative integer $-n_0$, all moments $\mu_n=0$ when $n>n_0$. In this case the formula for p_n defines polynomials only up to degree n_0. This degenerate case will not be considered.

Note further that the moments are not all positive when $\alpha<-1$. If $-j-1<\alpha<-j$, the first j moments alternate in sign. The remaining moments retain the same sign as μ_{j-1}.

Inserting the moments into the formula for w, we find

THEOREM 14.2. *For the Laguerre polynomials* $\{L_n^{(\alpha)}\}_{n=0}^{\infty}$,

$$w(x)=\sum_{n=0}^{\infty}(-1)^n\frac{\Gamma(n+\alpha+1)}{\Gamma(\alpha+1)}\frac{\delta^{(n)}(x)}{n!}$$

and

$$F^{-1}w(t)=(1/(2\pi))(1+it)^{-\alpha-1}.$$

When $\alpha>-1$, $F^{-1}w$ can easily be inverted by tables [6]:

THEOREM 14.3. *When* $\alpha>-1$,

$$w_P(x)=\begin{cases}\dfrac{x^{\alpha}e^{-x}}{\Gamma(\alpha+1)}, & x\geqq 0,\\ 0, & x<0.\end{cases}$$

Hence

$$\mu_n=\frac{\Gamma(n+\alpha+1)}{\Gamma(\alpha+1)}=\frac{1}{\Gamma(\alpha+1)}\int_0^{\infty}x^{n+\alpha}e^{-x}\,dx,$$

$n=0, 1, \cdots$, *and*

$$\langle w_P, L_m^{(\alpha)}L_n^{(\alpha)}\rangle=\frac{1}{\Gamma(\alpha+1)}\int_0^{\infty}L_m^{(\alpha)}L_n^{(\alpha)}(x)x^{\alpha}e^{-x}\,dx,$$

$$=\begin{cases}0 & \text{when } m\neq n,\\ \dfrac{n!\Gamma(n+\alpha+1)}{\Gamma(\alpha+1)} & \text{when } m=n.\end{cases}$$

A suitable weight function can also be found when $\alpha<-1$ and α is not a negative integer. Although $F^{-1}w$ cannot be directly inverted, a suitable derivative can be.

Let $-j-1<\alpha<-j$, and replace $(1+it)$ by z. Then

$$F^{-1}w_P(z)=z^{-\alpha-1}/(2\pi).$$

When $z=0$, $F^{-1}w_P=0$. Likewise when $0\leqq m<j$,

$$(F^{-1}w_P)^{(m)}(z)=(-1)^m(\alpha+1)\cdots(\alpha+m)z^{-\alpha-m-1}/(2\pi),$$

and $(F^{-1}w_P)^{(m)}|_{z=0}=0$. Finally,

$$(F^{-1}w_P)^j=(-1)^j(\alpha+1)\cdots(\alpha+j)z^{-\alpha-j-1}/(2\pi)$$

is the first derivative to become infinite at $z=0$, and is also the first to have a classical Fourier transform. Its transform is [6]

$$w_j(x)=\begin{cases}\dfrac{(-1)^j x^{\alpha+j}e^{-x}}{\Gamma(\alpha+1)}, & x\geqq 0,\\ 0, & x<0.\end{cases}$$

Hence

$$(F^{-1}w_P)^{(j)}(t)=\frac{(-1)^j}{2\pi\Gamma(\alpha+1)}\int_0^{\infty}x^{\alpha+j}e^{-x}e^{-itx}\,dx=\frac{(-1)^j}{2\pi\Gamma(\alpha+1)}\int_0^{\infty}x^{\alpha+j}e^{-xz}\,dx.$$

If integration in z is performed j times with limits from 0 to z, we find

$$F^{-1}w_P(z)=\frac{1}{2\pi\Gamma(\alpha+1)}\int_0^\infty x^\alpha\Big[e^{-xz}-\sum_{k=0}^{j-1}(-1)^k x^k z^k/k!\Big]\,dx$$
$$=\frac{1}{2\pi\Gamma(\alpha+1)}\int_0^\infty x^\alpha\Big[e^{-x}e^{-itx}-\sum_{k=0}^{j-1}(-1)^k x^k(1+it)^k/k!\Big]\,dx.$$

If the last term is expanded in powers of $-it$, and the summation indices are reversed, this becomes

$$F^{-1}w_P(t)=\frac{1}{2\pi\Gamma(\alpha+1)}\int_0^\infty x^\alpha\Big[e^{-x}e^{-itx}-\sum_{l=0}^{j-1}\Big\{\sum_{k=l}^{j-1}\frac{(-1)^k x^k}{l!(k-l)!}\Big\}(-1)^l(-it)^l\Big]\,dx.$$

Since $(-it)^l=(e^{-itx})^{(l)}$ evaluated at $x=0$, this suggests that the regularization of w is

$$\langle w_P,\psi\rangle=\frac{1}{\Gamma(\alpha+1)}\int_0^\infty x^\alpha\Big[e^{-x}\psi(x)-\sum_{l=0}^{j-1}\Big\{\sum_{k=l}^{j-1}\frac{(-1)^k x^k}{l!(k-l)!}\Big\}(-1)^l\psi^{(l)}(0)\Big]\,dx.$$

This agrees with [3]. An evaluation of $\langle w_P, x^n\rangle$ indeed verifies that

$$\langle w_P, x^n\rangle=\frac{\Gamma(n+\alpha+1)}{\Gamma(\alpha+1)},$$

$n=0,1,\cdots$. Thus

THEOREM 14.4. *Let $-j-1<\alpha<-j$. Then the Laguerre polynomials $\{L_n^{(\alpha)}(x)\}_{n=0}^\infty$ are mutually orthogonal with respect to the linear functional w_P defined by*

$$\langle w_P,\psi\rangle=\frac{1}{\Gamma(\alpha+1)}\int_0^\infty x^\alpha\Big[e^{-x}\psi(x)-\sum_{l=0}^{j-1}\Big\{\sum_{k=l}^{j-1}\frac{(-1)^k x^k}{l!(k-l)!}\Big\}(-1)^l\psi^{(l)}(0)\Big]\,dx.$$

Direct computation of the norms (squared) of $L_n^{(\alpha)}$, $\langle w, L_n^{(\alpha)2}\rangle$ is extremely awkward. By using the recurrence relation [7], however, they quickly follow:

THEOREM 14.5. *For all $\alpha\neq -1,-2,\cdots$,*

$$\langle w_P, L_n^{(2}\rangle=\frac{n!\Gamma(n+\alpha+1)}{\Gamma(\alpha+1)}$$

$n=0,1,\cdots$.

Proof. We multiply the relation

$$L_n^{(\alpha)}+(x-\alpha-2n-1)L_n^{(\alpha)}+n(n+\alpha)L_{n-1}^{(\alpha)}=0$$

by $L_{n+1}^{(\alpha)}$ and apply w_P to see

$$\langle w_P, L_{n+1}^{(\alpha)2}\rangle+\langle w_P, xL_n^{(\alpha)}L_{n+1}^{(\alpha)}\rangle=0.$$

If n is replaced by $n+1$ in the recurrence relation, it becomes

$$L_{n+2}^{(\alpha)}+(x-\alpha-2n-3)L_{n+1}^{(\alpha)}+(n+1)(n+\alpha+1)L_n^{(\alpha)}=0.$$

If this is multiplied by $L_n^{(\alpha)}$ and w_P is applied, then we have

$$\langle w_P, xL_{n+1}^{(\alpha)}L_n^{(\alpha)}\rangle+(n+1)(n+\alpha+1)\langle w_P, L_n^{(\alpha)2}\rangle=0.$$

Thus

$$\langle w_P, L_n^{(\alpha)2}\rangle=n(n+\alpha)\langle w_P, L_{n-1}^{(\alpha)2}\rangle.$$

The result follows by induction.

The values $n!\Gamma(n+\alpha+1)/\Gamma(\alpha+1)$, of course, coincide with previous results when $\alpha>-1$. When $\alpha<-1$, however, they are new. They oscillate in sign just as do the moments $\{\mu_i\}_{i=0}^{\infty}$.

15. The Jacobi polynomials. The Jacobi polynomials can also be made orthogonal in an extended sense, although a number of modifications in technique are required.

First the formulas of H. L. Krall [4] exhibited at the beginning of § 14 are as follows. If the differential equation for the Jacobi polynomials[2] $\{P_n\}_{n=0}^{\infty}$ is

$$(1-x^2)P_n''+[(u-v)-(u+v)x]P_n'+n(u+v+n-1)P_n=0,$$

the moment recurrence relation is

$$[u+v+n-1]\mu_n-[u+v]\mu_{n-1}-[n-1]\mu_{n-2}=0.$$

This is easily solved only in such simple cases as $u=0$, $v=0$ or $u=-v$, which are degenerate. Further, a direct computation of μ_n through the formula

$$\mu_n \cong \int_{-1}^{1} x^n(1-x)^{v-1}(1+x)^{u-1}\,dx$$

is available only when $u, v>0$. Consequently a direct computation of the moments does not seem reasonable.

Instead we follow a procedure developed by R. D. Morton. In the relations of H. L. Krall [4] replace x by $y=x-x_0$. Then the differential equation is transformed into

$$[l_{22}y^2+(2l_{22}x_0+l_{21})y+(l_{22}x_0^2+l_{21}x_0+l_{20})]p_n+[l_{11}y+(l_{11}x_0+l_{10})]p_n$$
$$=n[(l_{11}-l_{22})+l_{22}n]p_n,$$

and the recurrence relation becomes

$$(nl_{22}+l_{11}-l_{22})\mu_n(x_0)+([l_{11}+(n-1)l_{22}x_0+nl_{21}+l_{10}-l_{21})\mu_{n-1}(x_0)$$
$$+(n-1)(l_{22}x_0^2+l_{21}x_0+l_{20})\mu_{n-2}(x_0)=0,$$

where $\mu_n(x_0)$ is the nth moment about x_0.

If x_0 is chosen so that $l_{22}x_0^2+l_{21}x_0+l_{20}=0$, then the recurrence relation becomes "two term", and is easily solved.

For the Jacobi polynomials the recurrence relation is simplified when $-x_0^2+1=0$, or $x_0=\pm 1$.

THEOREM 15.1. *Let $\mu_0=1$. Then the Jacobi moments about* 1, $\mu_n(1)$, *are*

$$\mu_n(1)=\frac{(-1)^n 2^n (v)_n}{(u+v)_n},$$

$n=0, 1, \cdots$, *where* $(a)_n=a(a+1)\cdots(a+n-1)$.

Proof. The recurrence relation with $x_0=1$ is

$$\mu_n(1)(u+v+n-1)+\mu_{n-1}(1)2(v+n-1)=0,$$

which is solved by induction.

[2] Traditionally the Jacobi polynomials are indicated by $\{P_n^{\alpha,\beta}\}_{n=0}^{\infty}$. For notational purposes we set $\alpha=v-1$, $\beta=u-1$ and suppress α and β.

THEOREM 15.2. *Let* $\mu_0 = 1$. *Then the Jacobi moments about* -1, $\mu_n(-1)$, *are*

$$\mu_n(-1) = \frac{2^n (u)_n}{(u+v)_n},$$

$n = 0, 1, \cdots$.

As can be seen by inspection, there are three degenerate cases.

1. If $v = -N$, then $\mu_{N+1}(1) = 0$, and hence $\mu_n(1) = 0$ for all $n > N$. Only polynomials up to degree N exist.

2. If $u = -N$, then $\mu_{N+1}(-1) = 0$, and hence $\mu_n(-1) = 0$ for all $n > N$. Only polynomials up to degree N exist.

3. If $u + v = -N$, then either about 1 or -1 μ_{N+1} is undefined, as are the Jacobi polynomials.

We assume, therefore, that u, v, $u+v$ are not 0 or negative integers.

THEOREM 15.3. *Let* $\mu_0 = 1$. *Then the Jacobi moments about* 0, $\mu_n(0)$, *are*

$$\mu_n(0) = \sum_{j=0}^{n} \frac{\binom{n}{j}(-1)^j 2^j (v)_j}{(u+v)_j} = \sum_{j=0}^{n} \frac{\binom{n}{j}(-1)^{n-j} 2^j (u)_j}{(u+v)_j},$$

$n = 0, 1, \cdots$.

Proof.

$$\mu_n(0) = \langle w, x^n \rangle = \langle w, [(x-1)+1]^n \rangle$$

$$= \sum_{j=0}^{n} \binom{n}{j} \langle w, (x-1)^j \rangle = \sum_{j=0}^{n} \binom{n}{j} \mu_j(1).$$

Substitution yields the first expression. A similar expansion about -1 yields the second.

Since $\mu_0 = 1$, $\mu_1(1) = (u-v)/(u+v)$ in both expressions above, the recurrence relation verifies their equivalence for all n,

If we temporarily assume that u and v are complex with Re $u > 0$, Re $v > 0$, then the function

$$w_E(x) = \begin{cases} \dfrac{\Gamma(u+v)}{\Gamma(u)\Gamma(v)2^{u+v-1}}(1-x)^{v-1}(1+x)^{u-1}, & -1 \leqq x \leqq 1 \\ 0, & |x| > 1, \end{cases}$$

has as its inverse Fourier transform [6]

$$F^{-1} w_E(t) = e^{it} {}_1F_1(u, u+v, -2it)/(2\pi).$$

If this is expanded in a power series in $(-it)$, then

$$F^{-1} w_E(t) = \sum_{n=0}^{\infty} \left[\sum_{j=0}^{n} \frac{\binom{n}{j}(-1)^{n-j} 2^j (u)_j}{(u+v)_j} \right] \frac{(-it)^n}{n! 2\pi},$$

$$= \frac{1}{2\pi} \sum_{n=0}^{\infty} \mu_n(0)(-it)^n / n!,$$

where the moment formula for $\mu_n(0)$ using an expansion about -1 has been inserted. A comparison shows this formula agrees with the distributional inverse Fourier transform $F^{-1}w$.

Likewise, since

$$ {}_1F_1(a,b,z) = e^z\, {}_1F_1(b-a, b, -z), $$

$b \neq 0, -1, \cdots$, (see Rainville [7, p. 125]),

$$ \begin{aligned} F^{-1}w_E(t) &= e^{-it}\, {}_1F_1(v, u+v, 2it)/(2\pi), \\ &= \sum_{n=0}^{\infty} \left[\sum_{j=0}^{n} \binom{n}{j} \frac{(-1)^{n-j} 2^j (v)_j}{(u+v)_j} \right] \frac{(it)^n}{n!2\pi}, \\ &= \frac{1}{2\pi} \sum_{n=0}^{\infty} \mu_n(0)(-it)^n/n!, \end{aligned} $$

where the moment formula for $\mu_n(0)$ using an expansion about 1 has been inserted. A comparison again shows that $F^{-1}w_E$ agrees with $F^{-1}w$.

Rather than retrace the tedious calculations through a number of steps to find the various equivalent formulas when $\operatorname{Re} u < 0$, $\operatorname{Re} v < 0$ and $u, v, u+v \neq 0, -1, \cdots$, instead we use the principle of analytic continuation in u and v to achieve the results. We note that for $\operatorname{Re} u > 0$ and $\operatorname{Re} v > 0$, the following hold [7]:

Distributional formula.

$$ \langle w_E, \psi \rangle = \frac{\Gamma(u+v)}{\Gamma(u)\Gamma(v)2^{u+v-1}} \int_{-1}^{1} (1-x)^{v-1}(1+x)^{u-1}\psi(x)\, dx. $$

Inverse Fourier transform.

$$ F^{-1}w_E(t) = e^{it}\, {}_1F_1(u, u+v, -2it)/(2\pi) = e^{-it}\, {}_1F_1(v, u+v, 2it)/(2\pi). $$

Orthogonality.

$$ \langle w_E, P_n P_m \rangle = 0, \qquad n \neq m. $$

Norm squared.

$$ \langle w_E, P_n^2 \rangle = \frac{\Gamma(u+v)\Gamma(u+n)\Gamma(v+n)}{\Gamma(u)\Gamma(v)\Gamma(u+v+n-1)(u+v+2n-1)n!}, $$

THEOREM 15.4. *Let* $-M+1 > u > -M$, $-N+1 > v > -N$, $M, N > 0$. *Then the following hold*:

Distributional formula.

$$ \begin{aligned} \langle w_E, \psi \rangle = \left[\frac{\Gamma(u+v)}{\Gamma(u)\Gamma(v)2^{u+v-1}} \right] x \Bigg[& \int_0^1 (1-x)^{v-1} \Bigg\{ (1+x)^{u-1}\psi(x) \\ & - \sum_{j=0}^{N-1} \frac{[(1+x)^{u-1}\psi(x)]^{(j)}}{j!} \Bigg|_{x=1} (-1)^j (1-x)^j \Bigg\} dx \\ & + \int_{-1}^{0} (1+x)^{u-1} \Bigg\{ (1-x)^{v-1}\psi(x) - \sum_{k=0}^{M-1} \frac{[(1-x)^{v-1}\psi(x)]^{(k)}}{k!} \Bigg|_{x=-1} (1+x)^k \Bigg\} dx \\ & + \sum_{j=0}^{N-1} \frac{[(1+x)^{u-1}\psi(x)]^{(j)}}{j!} \Bigg|_{x=1} \frac{(-1)^j}{(v+j)} \\ & + \sum_{k=0}^{M-1} \frac{[(1-x)^{v-1}\psi(x)]^{(k)}}{k!} \Bigg|_{x=-1} \Bigg/ (u+k) \Bigg]. \end{aligned} $$

Inverse Fourier transform.

$$F^{-1}w_E(t)=e^{it}{}_1F_1(u,u+v,-2it)/(2\pi)=e^{-it}{}_1F_1(u,u+v,2it)/(2\pi).$$

Orthogonality.

$$\langle w_E, P_nP_m\rangle=0, \qquad n\neq m.$$

Norm squared.

$$\langle w_E, P_n^2\rangle=\frac{\Gamma(u+v)\Gamma(u+n)\Gamma(v+n)}{\Gamma(u)\Gamma(v)\Gamma(u+v+n-1)(u+v+2n-1)n!}.$$

Proof. It is clear that in each of these formulas the right sides are the analytic continuations of the right sides when Re $u>0$, Re $v>0$. The distributional formulas for the application of w_E to ψ, $e^{-ixt}/2\pi$, P_nP_m or P_n^2 are the canonical regularizations of the application of the analytic continuation of w_E, and in fact (see [3, p. 66]) the results are the analytic continuations of the integrals which result when Re $u>0$, Re $v>0$. As analytic continuations they must still agree.

16. The Bessel polynomials. Unfortunately the Bessel polynomials fail to yield completely to the techniques of this paper. The moments for the Bessel polynomials are $\mu_n=(-2)^{n+1}/(n+1)!$. Thus

$$w(x)=-\sum_{n=0}^{\infty}\frac{2^{n+1}\delta^{(n)}(x)}{n!(n+1)!}$$

and

$$F^{-1}w(t)=\frac{-1}{2\pi}\sum_{n=0}^{\infty}\frac{2^{n+1}(it)^n}{n!(n+1)!},$$

which is the power series representation for

$$f(t)=\frac{-1}{\pi}\frac{I_1((8it)^{1/2})}{(8it)^{1/2}}.$$

Since $|f(t)|\sim(1/\sqrt{2})\exp[2|t|^{1/2}]/(8\pi^2|t|)^{3/4}$ for large $|t|$, an extension of w beyond S still remains to be found.

Various tables (e.g. [8, #656.4]) show that $F^{-1}w$ corresponds to $(1/\pi)e^{-2/x}$. Direct calculation using the Bessel polynomial differential equation,

$$x^2y_n''+[2x+2]y_n'=n(n+1)y_n,$$

however, shows this is incorrect. The authors have devoted a great deal of time to extending w but are still left with the formula

$$w_P=F\left[\frac{-1}{\pi}\frac{I_1((8it)^{1/2})}{(8it)^{1/2}}\right],$$

which has defied evaluation.

It is possible to give a direct proof of orthogonality of the Bessel polynomials without resorting to the moment formula for p_n. All that is required is the existence of a weight function w or measure ν which gives the moments through the formulas $\mu_n=\langle w, x^n\rangle$, which certainly holds, or through $\mu_n=\int x^n\,d\nu$. Such a measure ν of bounded variation on $[0,\infty)$ is guaranteed by Boas [1].

The classical Bessel polynomials are given by the formula

$$y_n(x)=\sum_{k=0}^{n}\frac{(n+k)!}{(n-k)!k!}\left(\frac{x}{2}\right)^k,\qquad n=0,1,\cdots.$$

Since the highest order coefficient is $(2n)!/(2^n n!)$, $y_n=[(2n)!/(2^n n!)]p_n$.

Let us define

$$A_m^n=\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{(n+k)!}{(m+k+1)!}.$$

LEMMA 16.1. *Let* $0\leqq m<n$. *Then* $A_m^n=0$.

Proof. Consider

$$f(x)=x^n(1-x)^n=\sum_{k=0}^{n}(-1)^k\binom{n}{k}x^{n+k}.$$

Then

$$f^{(n-m-1)}(x)=\sum_{k=0}^{n}(-1)^k\binom{n}{k}\frac{(n+k)!}{(m+k+1)!}x^{m+k+1},$$

and $f^{(n-m-1)}(1)=A_m^n$.Since $f^{(n-m-1)}(1)=0$ when $0\leqq m<n$, so is A_m^n.

THEOREM 16.2. *The Bessel polynomials are mutually orthogonal with respect to w or with respect to an appropriate measure* ν *which generates the moments* $\mu_n=(-1)^{n+1}/(n+1)!$, $n=0,1,\cdots$.

Proof. If $0\leqq m<n$,

$$\begin{aligned}\langle w, y_n y_m\rangle&=\sum_{k=0}^{n}\sum_{j=0}^{m}\frac{(n+k)!(m+j)!\langle w, x^{j+k}\rangle}{(n-k)!k!(n-j)!j!2^{j+k}}\\&=2\sum_{k=0}^{n}\sum_{j=0}^{m}\frac{(n+k)!(m+j)!(-1)^{j+k+1}}{(n-k)!k!(m-j)!j!(j+k+1)!}\\&=2\sum_{j=0}^{m}\left[\frac{(m+j)!(-1)^{j+1}}{(m-j)!j!n!}\right]A_j^n=0.\end{aligned}$$

The norms (squared) can also be readily calculated. Define

$$B_m^n=2\sum_{k=0}^{n}\sum_{j=0}^{m}\frac{(n+k)!(n+j)!(-1)^{j+k+1}}{(n-k)!k!(m-j)!j!(j+k+1)!}.$$

LEMMA 16.3. *If* $m=n\geqq 1$, *then* $B_n^n=(-1)^{n+1}2/(2n+1)$.

Proof. Again consider

$$f(x)=x^n(1-x)^n=\sum_{k=0}^{n}(-1)^k\binom{n}{k}x^{n+k}.$$

Let

$$F(y)=\int_0^y f(x)\,dx=\sum_{k=0}^{n}(-1)^k\binom{n}{k}y^{n+k+1}\Big/(n+k+1).$$

Then $F(1)=A_n^n$. But if integration by parts is performed n times,

$$F(1)=[(n!)^2/(2n)!]\int_0^1(1-x)^{2n}\,dx=(n!)^2/(2n+1)!.$$

Thus $A_n^n = (n!)^2/(2n+1)!$ and

$$B_n^n = 2[(2n)!!(-1)^{n+1}/(n!)^2]A_n^n = (-1)^{n+1}2/(2n+1).$$

We have only to note that either $\langle w, y_n^2\rangle$ or $\int y_n^2\, d\nu$ equal B_n^n to conclude

THEOREM 16.4. *The Bessel polynomials satisfy*

$$\left.\begin{matrix} \langle w, y_n^2\rangle \\ \textit{or} \\ \int y_n^2\, d\nu \end{matrix}\right\} = (-1)^n 2/(2n+1).$$

This is in agreement with the calculation done by H. L. Krall and O. Frink [5] using a different method.

REFERENCES

[1] R. P. BOAS, JR., *The Stieltjes moment problem for functions of bounded variation*, Bull. Amer. Math. Soc., 45 (1939), pp. 399–404.

[2] H. BREMERMANN, *Distributions, Complex Variables, and Fourier Transforms*, Addison-Wesley, Reading, MA, 1965.

[3] I. M. GEL'FAND AND G. E. SHILOV, *Generalized Functions*, vol. I, Academic Press, New York, 1964.

[4] H. L. KRALL, *Certain differential equations for Tchebycheff polynomials*, Duke Math. J., 4 (1938), pp. 705–718.

[5] H. L. KRALL AND O. FRINK, *A new class of orthogonal polynomials: The Bessel polynomials*, Trans. Amer. Math. Soc., 65 (1949), pp. 100–115.

[6] F. OBERHETTINGER, *Tabellen zur Fourier Transformen*, Springer-Verlag, Berlin, 1957.

[7] E. D. RAINVILLE, *Special Functions*, Macmillan, New York, 1960.

[8] B. W. ROOS, *Analytic Functions and Distributions in Physics and Engineering*, John Wiley, New York, 1969.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0006 $01.00/0

AN ADDITION THEOREM FOR HAHN POLYNOMIALS: THE SPHERICAL FUNCTIONS*

CHARLES F. DUNKL†

Abstract. An addition formula for the Hahn polynomials $Q_k(x;\alpha,\beta,N)$ is derived for the parameter values $\beta=-N-1$, $\alpha\neq -1,-2,\cdots,-N$, $N=1,2,3,\cdots$. The method is to realize Q_k as a spherical function for the values $\alpha=-N-1,-N-2,\cdots$ and to use harmonic analysis on the finite homogeneous space $(S_b\times S_a)\backslash S_{a+b}$ where $b=N$, $a=-\alpha-1$ and S_n is the symmetric group on n objects ($n=1,2,\cdots$).

Introduction. The Hahn polynomials form a three-parameter family of orthogonal polynomials; for an integer $N\geqq 1$, and real numbers α,β the corresponding weight function is $(\alpha+1)_x(\beta+1)_{N-x}/(x!(N-x)!)$ at $x=0,1,\cdots,N$, of constant sign and thus providing actual orthogonality, for the values $\alpha,\beta>-1$ or $\alpha,\beta<-N$. The polynomials are denoted by $Q_k(x;\alpha,\beta,N)$, and $Q_k(x;\alpha,\beta,N)$

$$={}_3F_2\left(\begin{matrix}-k,k+\alpha+\beta+1,-x\\-N,\alpha+1\end{matrix};1\right)=\sum_{i=0}^{k}\frac{(-k)_i(k+\alpha+\beta+1)_i(-x)_i}{(-N)_i(\alpha+1)_i i!}$$

by definition, $k=0,1,\cdots,N$. The Pochhammer symbol $(a)_i$ is defined by $(a)_0=1$, $(a)_i=a(a+1)\cdots(a+i-1)$, $i=1,2,\cdots$.

This paper contains the addition formula and its proof, for the parameter values $\beta=-b-1$, $N=b$, $b=1,2,\cdots$, $\alpha\neq-1,-2,\cdots,-b$. For reasons of symmetry we use the parameter $a=-\alpha-1$. We define an auxiliary function E_m, $m=0,1,2,\cdots$ by

$$E_m(a,b,c,x)=\sum_{j=0}^{m}(-1)^j\binom{m}{j}(b-m+1)_j(-x)_j(a-m+1)_{m-j}(x-c)_{m-j}.$$

The addition formula is:

$$\begin{aligned}Q_k(v+w-x-y;-a-1,-b-1,b)=\sum_m\sum_n c_{mnk}(a,b)\\ \cdot E_{k-m-n}(a-2n,b-2m,b-m-n,v-n)\\ \cdot E_{k-m-n}(a-2n,b-2m,b-m-n,w-n)\\ \cdot E_m(w,b-w,v,x)E_n(w,a-w,v,y),\end{aligned}$$

where

$$c_{mnk}(a,b)=\frac{(-1)^{m+n}(-k)_{m+n}(k-a-b-1)_{m+n}(b-2m+1)(a-2n+1)}{n!m!(-a)_k(-b)_k(n-a)_{k-m}(m-b)_{k-n}(b-m+1)(a-n+1)}.$$

a,b,k are integers with $0\leqq k\leqq b\leqq a$, the sum is taken over integers m,n with $0\leqq m+n\leqq k$, $0\leqq m\leqq b/2$, $0\leqq n\leqq a/2$, $k-a\leqq m-n\leqq b-k$, and v,w,x,y are integers such that $0\leqq w$, $v\leqq b$, $\max(0,w+v-b)\leqq x\leqq\min(w,v)$, $\max(0,w+v-a)\leqq y\leqq\min(w,v)$. For nonintegral values of a see Corollary 5.3.

* Received by the editors August 3, 1976, and in revised form November 9, 1976.

† Department of Mathematics, University of Virginia, Charlottesville, Virginia 22903. This work was supported in part by the National Science Foundation under Grant MCS 76-07022.

1. Outline and notation. The derivation of the addition formula depends on the fact that the functions $Q_k(\cdot\,; -a-1, -b-1, b)$ (integers a, b with $a \geqq b \geqq 1$) are the spherical functions for the homogeneous space $X = (S_b \times S_a)\backslash S_{a+b}$ where S_n is the symmetric group on n objects ($n = 1, 2, \cdots$). This space is realized as the space of b-subsets of a fixed $(a+b)$-set. The addition formula for a spherical function is essentially the Fourier series with respect to the stabilizer subgroup, $S_b \times S_a$, of an arbitrary right translate of the function. Our method depends on each term of the Fourier series satisfying a difference equation and being invariant under a certain subgroup. The subsequent part of this paper is divided as follows:

2. $L(X)$ and its decomposition into G-modules: the S_{a+b}-invariant subspaces V_k of the functions on X, described by difference and formal differential operators; the spherical functions.

3. The auxiliary functions E_m: relations to ${}_3F_2$ series, Hahn polynomials, symmetries, orthogonality, difference relations.

4. Splitting V_k into H-modules: the further decomposition into subspaces invariant under $S_b \times S_a$, characterization of functions invariant under certain subgroups.

5. The addition theorem: application of harmonic analysis to put the addition formula together out of the previously constructed ingredients.

1.1. Notation. For x, y real, we will use $x \wedge y$, $x \vee y$ for $\min(x, y)$, $\max(x, y)$ respectively. For integers c, d with $c \leqq d$, $[c, d]$ means the set $\{c, c+1, \cdots, d\}$. For a set η, $|\eta|$ denotes the cardinality.

Fix integers a, b with $1 \leqq b \leqq a$. Let $G = S_{a+b}$, the symmetric group acting on $[1, a+b]$ (permutation written on right side of point). For $\eta \subset [1, a+b]$, let $S(\eta)$ be the subgroup of G which fixes each point in $[1, a+b]\backslash\eta$ (essentially $S(\eta)$ is the permutation group of η). Let $H = S([1, b]) \times S([b+1, a+b])$. For $w = 0, 1, \cdots, b$ define K_w as follows: $K_0 = H$ and for $1 \leqq w \leqq b$, $K_w = S([1, w]) \times S([w+1, b]) \times S([b+1, a+b-w]) \times S([a+b-w+1, a+b]) \subset H$.

We will be concerned with the homogeneous space $H\backslash G$ which will be realized as $\{\eta : \eta \subset [1, a+b], |\eta| = b\}$ and denoted by X; note $|X| = \binom{a+b}{b}$. Let $L(X)$ be the space of complex functions on X furnished with the inner product $\langle f_1, f_2 \rangle = (1/|X|) \sum_{\xi \in X} f_1(\xi)\overline{f_2(\xi)}$ and norm $\|f_1\| = \langle f_1, f_1 \rangle^{1/2} (f_1, f_2 \in L(X))$. The representation of G by right translation on $L(X)$ will be denoted by R, thus $\mathrm{R}(g)f(\xi) = f(\xi g)(f \in L(X), \xi \in X, g \in G)$. There is a G-invariant metric ρ on X, namely $\rho(\xi, \eta) = b - |\xi \cap \eta|$. In coding theory terminology, X is called a Johnson association scheme with the Lee metric ρ. For more details and bibliography on this, as well as an association-scheme-theoretic derivation of the role of Hahn polynomials, see Delsarte [2], [3].

It is sometimes convenient to use polynomial functions on X, thus we define coordinate functions x_i, $i = 1, \cdots, a+b$ by $x_i(\xi) = 1$ if $i \in \xi$, 0 if $i \notin \xi$ for $\xi \subset [1, a+b]$; thus X is embedded as a subset of $\mathbb{R}^{a+b}$, and G is represented as a linear group on $\mathbb{R}^{a+b}$, acting by permutation of coordinates. The base point in X, namely $[1, b]$, will be denoted ω; of course H is the stabilizer of ω.

2. $L(X)$ and its decomposition into G-modules. For each $\eta \subset [1, a+b]$ let x_η denote $\prod_{i \in \eta} x_i$ (1 for η empty). For $0 \leqq m \leqq a+b$ let P_m be the linear span of $\{x_\eta : |\eta| = m\}$. Then $\dim P_m = \binom{a+b}{m}$ and P_m is invariant under right translation by G, but is not in general irreducible. Let $d = \sum_{i=1}^{a+b} \partial/\partial x_i$, a formal differential operator which commutes with $R(g)(g \in G)$, and define $V_m = P_m \cap \ker d$, $0 \leqq m \leqq b$. It was shown in [4, pp. 342–344] that d maps P_m onto P_{m-1} for $b \geqq m \geqq 1$, that V_m is irreducible (under G) of dimension

$$\binom{a+b}{m} - \binom{a+b}{m-1} = \binom{a+b}{m}\left(\frac{a+b-2m+1}{a+b-m+1}\right),$$

and realizes the representation $[a+b-m, m]$ of G. Thus the V_m's are pairwise orthogonal and $L(X) = \sum_{m=0}^{b} \oplus V_m$.

We now find polynomials with specified invariance properties. In particular, fix disjoint subsets $\eta_1, \eta_2 \subset [1, a+b]$. We look for elements of V_m which are $S(\eta_1) \times S(\eta_2)$-invariant and involve only the variables x_i, $i \in \eta_1 \cup \eta_2$. Indeed these are linear combinations of $\sigma_i(\eta_1)\sigma_{m-i}(\eta_2)$ $(0 \leqq i \leqq m)$, where $\sigma_j(\eta)$ is the elementary symmetric function of degree j in the variables x_i, $i \in \eta \subset [1, a+b]$. The requirement of being in the kernel of d determines these polynomials uniquely (up to a multiplicative constant), see [4, 2.12, p. 345]; the key fact is that $d\sigma_j(\eta) = (|\eta| - j + 1)\sigma_{j-1}(\eta)$.

THEOREM 2.1. *Let η_1, η_2 be disjoint subsets of $[1, a+b]$ and let $p \in V_m$ $(0 \leqq m \leqq b)$ be $S(\eta_1) \times S(\eta_2)$-invariant and involve only the variables x_i, $i \in \eta_1 \cup \eta_2$. Then*

$$p = c \sum_{i=0}^{m} (|\eta_2| - m + 1)_i (|\eta_1| - m + 1)_{m-i} (-1)^i \sigma_i(\eta_1)\sigma_{m-i}(\eta_2)$$

(some $c \in \mathcal{C}$) if $m \leqq |\eta_1| \wedge |\eta_2|$, and $p \equiv 0$ otherwise, although the formula holds for all m $(\sigma_i(\eta) \equiv 0$ if $i > |\eta|)$.

The theorem applied to $\eta_1 = [1, b]$, $\eta_2 = [b+1, a+b]$ gives the spherical functions (see [4, p. 352]).

COROLLARY 2.2. *The spherical function for V_m, namely, the unique element ϕ_m of V_m such that ϕ_m is H-invariant and has $\phi_m(\omega) = 1$, is $\phi_m(\xi) = Q_m(v(\xi); -a-1, -b-1, b)$ $(\xi \in X)$ where $v(\xi) = b - |[1, b] \cap \xi| = \rho(\omega, \xi)$.*

We will need a difference equation criterion for membership in V_m. Any operator on $L(X)$ which commutes with $R(g)(g \in G)$ will have the V_m's as eigenmanifolds, so we set up a "locally-defined" example of such an operator. Indeed, define

$$Tf(\xi) = \sum \{f(\zeta) : \zeta \in X, |\zeta \cap \xi| = b-1\} \qquad (f \in L(X), \xi \in X),$$

(that is, sum the values over the adjacent points in X, $\rho(\xi, \zeta) = 1$).

THEOREM 2.3. *Let $f \in L(X)$, $0 \leqq m \leqq b$; then $f \in V_m$ if and only if $Tf = [ab - m(a+b+1-m)]f$.*

Proof. Consider $f \in L(X)$ as being a left H-invariant function on G (mapping $g \mapsto f(\omega g)$, $g \in G$), and let k be a zonal (H-H invariant) function so that $k(g) = k_1(v(\omega g))$ (see Corollary 2.2). Then $k * f$ is left H-invariant on G, thus an element of $L(X)$ (indeed every G-operator on $L(X)$ is of the form $f \mapsto k * f$). Note that $k * f(g) = (1/|G|)\sum_{g_1} k(gg_1^{-1})f(\omega g_1) = (1/|G|)\sum_{g_1} k_1(\rho(\omega g, \omega g_1))f(\omega g_1) = (1/|X|)\sum_{\xi \in X} k_1(\rho(\omega g, \xi))f(\xi)$. Now define k by $k_1(1) = |X|$, $k_1(j) = 0$, $j \neq 1$; then $k * f = Tf$. On the other hand, write $f = \sum_{m=0}^{b} f_m$, the Fourier series of f, where $f_m \in V_m$; then $k * f = \sum_{m=0}^{b} \hat{k}_m f_m$, where $\hat{k}_m = (1/|G|)\sum_{g \in G} k(g)\phi_m(\omega g) = (|X|/|G|)|\{g : \rho(\omega, \omega g) = 1\}| Q_m(1; -a-1, -b-1, b) = ab(1 - m(a+b+1-m)/ab)$. The eigenvalues determine m uniquely in the range $0 \leqq m \leqq b$ $(\leqq (a+b+1)/2)$, thus the theorem is proved. □

3. The auxiliary functions E_m. Let m, a, b, c, x be nonnegative integers and define

$$E_m(a, b, c, x) = \sum_{j=0}^{m} (-1)^j \binom{m}{j} (b-m+1)_j (-x)_j (a-m+1)_{m-j} (x-c)_{m-j},$$

a polynomial of degree m in each of the variables a, b, c, x. Of course, E_m can be directly written as a ${}_3F_2$, indeed

$$E_m(a, b, c, x) = (-1)^m(-a)_m(x-c)_m \, {}_3F_2\left(\begin{matrix} b-m+1, -x, -m \\ -a, c-x-m+1 \end{matrix}; 1\right), \tag{3.1}$$

but we will use the E_m notation for three main reasons: 1) E_m directly shows polynomial dependence on variables without the problem of canceling out possible zero factors in denominators, 2) certain symmetries are much more neatly expressible, 3) the parameters a, b, c specify appropriate constraints for the domain of the x variable (and 4) protect the author and the reader from being case-wised to the point of loathing).

We will exhibit the relation of E_m to the Hahn polynomials, leading to the orthogonality relations, and also find some difference equations satisfied by the E_m's, various symmetries and special values.

First, interchanging j and $m-j$ shows:

$$E_m(a, b, c, x) = (-1)^m E_m(b, a, c, c-x). \tag{3.2}$$

Next we get a deeper identity:

For $a, c \neq 0, 1, \cdots, m-1$,

$$E_m(a, b, c, x) = (-1)^m(-a)_m(-c)_m \, {}_3F_2\left(\begin{matrix} m-a-b-1, -x, -m \\ -a, -c \end{matrix}; 1\right). \tag{3.3}$$

This follows from (3.1) by using the following transformation of a terminating ${}_3F_2$, $m = 1, 2, \cdots$:

$${}_3F_2\left(\begin{matrix} -m, a, b \\ c, d \end{matrix}; 1\right) = \left(\frac{(c-b)_m}{(c)_m}\right) {}_3F_2\left(\begin{matrix} -m, d-a, b \\ b-c-m+1, d \end{matrix}; 1\right).$$

One way to prove this is to begin with a formula of Pfaff,

$${}_2F_1\left(\begin{matrix} -m, b \\ c \end{matrix}; x\right) = \left(\frac{(c-b)_m}{(c)_m}\right) {}_2F_1\left(\begin{matrix} -m, b \\ b-c-m+1 \end{matrix}; 1-x\right)$$

(which comes from expanding $(1-(1-x))^j$ by the binomial theorem and using the Chu–Vandermonde sum), multiply both sides by $x^{a-1}(1-x)^{d-a-1}$ and integrate over $0 \leqq x \leqq 1$. (The author thanks R. Askey for this suggestion.) The transformation can also be found in Bailey's book [1, p. 22], and was used by Gasper [6, 2.6] on Hahn polynomials.

As a corollary of (3.3) we obtain:

$$E_m(a, b, c, x) = E_m(c, a+b-c, a, x) \quad \text{valid for all } a, b, c, x. \tag{3.4}$$

Combining (3.2) and (3.4) several times we get

$$\begin{aligned} E_m(a, b, c, x) &= (-1)^m E_m(a+b-c, c, a, a-x) && \text{(by 3.2)} \\ &= (-1)^m E_m(a, b, a+b-c, a-x) && \text{(by 3.4)} \\ &= E_m(b, a, a+b-c, x+b-c) && \text{(by 3.2).} \end{aligned} \tag{3.5}$$

To get orthogonality relations and nontrivial Hahn polynomials we now impose the conditions

$$0 \leqq c \leqq a+b, \quad 2m \leqq a+b, \quad (c-b) \vee 0 \leqq x \leqq a \wedge c.$$

The E_m's are expressed as Hahn polynomials in four different ways depending on a, b, c.

$$\text{(i)}\quad c \leqq b,\ c \leqq a,\quad E_m(a, b, c, x) = (-1)^m(-a)_m(-c)_m Q_m(x; -a-1, -b-1, c), \quad \text{(by 3.3)}$$

$$\text{(ii)}\quad a \leqq c \leqq b,\quad E_m(a, b, c, x) = (-1)^m(-a)_m(-c)_m Q_m(x; -c-1, -(a+b-c)-1, a),$$

$$\text{(3.6)}\quad \text{(iii)}\quad b \leqq c \leqq a,\quad E_m(a, b, c, x) = (-1)^m(-b)_m(c-a-b)_m Q_m(x+b-c; -b-1, -a-1, a+b-c), \quad \text{(by 3.5 and (i))}$$

$$\text{(iv)}\quad a \leqq c,\ b \leqq c,\quad E_m(a, b, c, x) = (-1)^m(-b)_m(c-a-b)_m Q_m(x+b-c, -(a+b-c)-1, -c-1, b).$$

By using the orthogonality relations of the Q_m's (see Karlin and McGregor [7] or [4, § 5]) we find

$$\text{(3.7)}\quad \sum_{x=0\vee(c-b)}^{a\wedge c} \binom{a}{x}\binom{b}{c-x} E_m(a, b, c, x)E_n(a, b, c, x) = \delta_{mn}\binom{a+b}{c}\binom{a+b}{m}^{-1}\left(\frac{a+b-m+1}{a+b-2m+1}\right) \cdot (-a)_m(-b)_m(-c)_m(c-a-b)_m, \qquad 0 \leqq m, n \leqq (a+b)/2.$$

If $m > a\wedge b\wedge c\wedge(a+b-c)$, then $E_m(a, b, c, x) = 0$ for (integral) x in the domain $(c-b)\vee 0 \leqq x \leqq a\wedge c$. There are four special values of E_m, namely

$$\text{(3.8)}\quad \begin{aligned} &\text{(i)} && E_m(a, b, c, 0) = (-1)^m(-a)_m(-c)_m && \text{(by 3.1)},\\ &\text{(ii)} && E_m(a, b, c, c) = (-b)_m(-c)_m && \text{(by 3.2 and (i))},\\ &\text{(iii)} && E_m(a, b, c, a) = (c-a-b)_m(-a)_m && \text{(by 3.5)},\\ &\text{(iv)} && E_m(a, b, c, c-b) = (-1)^m(c-a-b)_m(-b)_m && \text{(by 3.5)}, \end{aligned}$$

(note that (iv) is the Saalschütz formula for (3.3)). The E_m's satisfy certain difference relations:

$$\text{(3.9)}\quad (c-x)E_m(a, b, c-1, x) + xE_m(a, b, c-1, x-1) = (c-m)E_m(a, b, c, x),$$

$$\text{(3.10)}\quad (a-x)E_m(a, b, c+1, x+1) + (b-c+x)E_m(a, b, c+1, x) = (a+b-c-m)E_m(a, b, c, x).$$

It is easy to prove (3.9) by using $(c-x)(-x)_j(x-c+1)_{m-j} + x(-x+1)_j(x-c)_{m-j} = (c-m)(-x)_j(x-c)_{m-j}$. Interpreted for Hahn polynomials by (3.6(i)), it is a special case of a formula of Gasper [6, 2.3]. To get (3.10) transform $E_m(a, b, c+1, x)$ to $E_m(b, a, a+b-c-1, x+b-c-1)$ by (3.5) and use (3.9). Now replace c by $c+1$ in (3.9), apply to both x and $x+1$, and then substitute in the left hand side of (3.10) to obtain

$$\text{(3.11)}\quad (c-x)(a-x)E_m(a, b, c, x+1) + x(x+b-c)E_m(a, b, c, x-1) = [(c-x)(a-x) + x(x+b-c) - m(a+b+1-m)]E_m(a, b, c, x).$$

This identity was discovered by Karlin and McGregor [7, 1.3], in Hahn polynomial terms.

The difference equation eigenvalue problem corresponding to (3.11) is

$$B(x)f(x+1)+D(x)f(x-1)-(B(x)+D(x))f(x)=\lambda f(x),$$
$$B(x)=(c-x)(a-x), \quad D(x)=x(x+b-c). \tag{3.12}$$

Notice the singular points are 0, $c-b$, c, a (compare with (3.8)). For a given value $f(x_0)$, where $x_0=0\vee(c-b)$, the value of λ and (3.12) determine the values of $f(x_0+1),\cdots,f(a\wedge c)$ (with $x=x_0, x_0+1,\cdots,a\wedge c-1$), and the equation for $x=a\wedge c$ is a polynomial of degree $(a\wedge c)-x_0$ in λ, with zeros $-m(a+b+1-m)$, $m=0,1,\cdots a\wedge b\wedge c\wedge(a+b-c)$. This is of course clear; the point is to emphasize that the E_m's are the only solutions of (3.12).

4. Splitting V_k into H-modules. Recall from § 2 that V_k realizes the irreducible representation $[a+b-k,k]$ of G. From the representation theory of the symmetric group (see Robinson [9]) we find that the restriction of $[a+b-k,k]$ ($k\leqq(a+b)/2$) to H splits as follows:

$$[a+b-k,k]\|H\cong\sum\oplus\{[b-m,m]\otimes[a-n,n]: 0\leqq m\leqq b/2,$$
$$0\leqq n\leqq a/2, m+n\leqq k, k-a\leqq m-n\leqq b-k\}. \tag{4.1}$$

DEFINITION 4.1. For integers m, n, $0\leqq m\leqq b/2$, $0\leqq n\leqq a/2$, let P_{mn} be the orthogonal projection of V_k onto the subspace giving the representation $[b-m,m]\otimes[a-n,n]$ of H (if V_k does not contain $[b-m,m]\otimes[a-n,n]$ set $P_{mn}=0$). Each P_{mn} commutes with $R(h)(h\in H)$.

We now find functions in $P_{mn}V_k$ which are invariant under K_w, for some $w\in[0,b]$ (see § 1.1). Note that the K_w-invariant functions are those which are constant on the K_w-orbits in X. The K_w-orbit of $\zeta\in X$ is determined by the four coordinates $y_i(\zeta)=|\zeta\cap\eta_i|$, $i\in[1,4]$, where $\eta_1=[1,w]$, $\eta_2=[w+1,b]$, $\eta_3=[b+1,a+b-w]$, $\eta_4=[a+b-w+1,a+b]$ (one y_i is redundant since $\sum_i y_i=b$).

The space $P_{mn}V_k$ is easy to describe when $k=m+n$ (and $0\leqq m\leqq b/2$, $0\leqq n\leqq a/2$). We need merely form the spaces V_m^b, V_n^a in the variables $x_1,\cdots,x_b$ and $x_{b+1},\cdots,x_{a+b}$ respectively, so V_m^b is the space of square-free (no x_i^2) polynomials in $x_1,\cdots,x_b$, homogeneous of degree m, satisfying $\sum_{i=1}^b \partial p/\partial x_i=0$, and V_n^a is similarly defined (in the variables $x_{b+1},\cdots,x_{a+b}$). The product mapping $(p,q)\mapsto pq$ extends to a linear map of $V_m^b\otimes V_n^a$ into V_{m+n}, indeed onto $P_{mn}V_{m+n}$. By the product rule for differentiation the polynomial pq satisfies $d(pq)=0$ ($p\in V_m^b$, $q\in V_n^a$).

LEMMA 4.2. *For $0\leqq m\leqq b/2$, $0\leqq n\leqq a/2$, $P_{mn}V_{m+n}$ is the linear span of $\{pq: p\in V_m^b, q\in V_n^a\}$*

LEMMA 4.3. *Let $0\leqq m\leqq b/2$, $0\leqq n\leqq a/2$, $0\leqq w\leqq b$, $f\in P_{mn}V_{m+n}$ and let f be K_w-invariant; then f is a scalar multiple of $g_m^b g_n^a$, where $g_m^b(y_1,y_2)=E_m(w,b-w,y_1+y_2,y_1)$ and $g_n^a(y_3,y_4)=E_n(w,a-w,y_3+y_4,y_4)$ (recall the y_i coordinates, $y_i(\zeta)=|\zeta\cap\eta_i|$, $i\in[1,4]$).*

Proof. Applying Theorem 2.1 to $S(\eta_1)\times S(\eta_2)$ and $S(\eta_3)\times S(\eta_4)$ shows that the unique K_w-invariant elements of V_m^b, V_n^a are $\sum_{i=0}^m(-1)^i(|\eta_2|-m+1)_i(|\eta_1|-m+1)_{m-i}\sigma_i(\eta_1)\sigma_{m-i}(\eta_2)$ and $\sum_{j=0}^n(-1)^j(|\eta_3|-n+1)_j(|\eta_4|-n+1)_{n-j}\sigma_j(\eta_4)\sigma_{n-j}(\eta_3)$ respectively. Evaluating $\sigma_i(\eta_1)\sigma_{m-i}(\eta_2)$ and $\sigma_j(\eta_4)\sigma_{n-j}(\eta_3)$ at (y_i) yields $\binom{y_1}{i}\binom{y_2}{m-i}$ and $\binom{y_4}{j}\binom{y_3}{n-j}$, that is, $(-1)^m(-y_1)_i(-y_2)_{m-i}/(i!(m-i)!)$ and $(-1)^n(-y_4)_j(-y_3)_{n-j}/(j!(n-$

$j)!$) respectively, and so we get the functions E_m, E_n as indicated (up to multiplicative constants). The product of these two functions is K_w-invariant, and in $P_{mn}V_{m+n}$ by Lemma 4.2, and is the only such function (up to scalar multiplication). □

Suppose that f is H-invariant on X; then for each $\zeta \in X$ the function $h \mapsto f(\zeta h)(g^b_m g^a_n)(\zeta h)$, $(h \in H)$, is a trigonometric polynomial from $[b-m, m] \otimes [a-n, n]$ and is K_w-invariant. We know that $P_{mn}V_k$ contains a unique K_w-invariant element (because it is isomorphic as an H-module to $P_{mn}V_{m+n}$). If we can find f so that $fg^b_m g^a_n \in V_k$, then $fg^b_m g^a_n \in P_{mn}V_k$ and is the K_w-invariant element of this H-module. Indeed we can do this by using Theorem 2.3, for which purpose we must express T in the coordinates y_i. Let $c_i = |\eta_i|$, $i \in [1, 4]$.

LEMMA 4.4. *Let f be a K_w-invariant function on X, given in terms of the y_i coordinates, then* $Tf(\zeta) = \sum_{i=1}^4 y_i(c_i - y_i)f(\zeta) + \sum_{i \neq j} y_j(c_i - y_i)f(\cdots, y_i + 1, \cdots, y_j - 1, \cdots)$ *(implying the other coordinates are fixed, and i not necessarily $<j$).*

Proof. Fix $\zeta \in X$. Adjacent points are to be counted according to the y_i coordinates. For each i, interchanging a point in $\zeta \cap \eta_i$ (y_i such) with a point in $\eta_i \backslash \zeta$ ($c_i - y_i$ such) produces a point in X adjacent to ζ with the same y_i coordinates. For each ordered pair $i, j, i \neq j$, interchanging a point in $\zeta \cap \eta_j$ (y_j such) with one in $\eta_i \backslash \zeta$ ($c_i - y_i$ such), produces the coordinates $y_i + 1$, $y_j - 1$, other y's the same. □

To get coordinates for K_w-orbits compatible with the description of H-orbits we introduce the following: $v = y_3 + y_4$ (agrees with notation of Corollary 2.2), $x = y_1$, $y = y_4$ (thus $y_2 = b - v - x$, $y_3 = v - y$). To describe T in terms of the v, x, y coordinates we present a list of the thirteen possibilities in terms of (y_i) and their corresponding v, x, y values (for brevity only the changed y_i values will appear):

Case 1: same v, 1) same x, y, 2) $y_1 + 1, y_2 - 1, x + 1, y$, 3) $y_1 - 1, y_2 + 1, x - 1, y$, 4) $y_3 - 1, y_4 + 1, x, y + 1$, 5) $y_3 + 1, y_4 - 1, x, y - 1$;

Case 2: $v + 1$, one of (a) $y_1 - 1, x - 1$, (b) $y_2 - 1, x$ with one of (c) $y_3 + 1, y$, (d) $y_4 + 1, y + 1$;

Case 3: $v - 1$, one of (a) $y_1 + 1, x + 1$, (b) $y_2 + 1, x$ with one of (c) $y_3 - 1, y$, (d) $y_4 - 1, y - 1$.

Recall $c_i = |\eta_i|$, $i \in [1, 4]$, so $c_1 = c_4 = w$, $c_2 = b - w$, $c_3 = a - w$. In the expression for $T(fg^b_m g^a_n)$ from Lemma 4.4 some simplifications in the way of products are possible, and the following is obtained:

$$\begin{aligned} T(fg^b_m g^a_n)(x, y, v) = f(v)[&p_0(v)g^b_m(x, b-v-x)g^a_n(v-y, y) \\ &+ p_{10}(x, v)g^a_n(v-y, y) + p_{01}(y, v)g^b_m(x, b-v-x)] \\ &+ f(v+1)p_{20}(x, v)p_{02}(y, v) + f(v-1)p_{30}(x, v)p_{03}(y, v), \end{aligned}$$

where

$$\begin{aligned} p_0(v) &= x(w-x) + (b-v-x)(x+v-w) + (v-y)(a-w-v+y) + y(w-y), \\ p_{10}(x, v) &= (b-v-x)(w-x)E_m(w, b-w, b-v, x+1) \\ &\qquad + x(x+v-w)E_m(w, b-w, b-v, x-1) \\ &= [(b-v-x)(w-x) + x(x+v-w) - m(b+1-m)]E_m(w, b-w, v, x) \\ &\qquad \text{(by 3.11)}, \\ p_{01}(y, v) &= (v-y)(w-y)E_n(w, a-w, v, \hat{y}+1) + y(a-w-v+y)E_n(w, a-w, v, y-1) \\ &= [(v-y)(w-y) + y(a-w-v+y) - n(a+1-n)]E_n(w, a-w, v, y) \\ &\qquad \text{(by 3.11)}, \end{aligned}$$

$$\begin{aligned} p_{20}(x, v) &= xg_m^b(x-1, b-v-x) + (b-v-x)g_m^b(x, b-v-1-x) \\ &= xE_m(w, b-w, b-v-1, x-1) + (b-v-x)E_m(w, b-w, b-v-1, x) \\ &= (b-v-m)E_m(w, b-w, b-v, x) \qquad \text{(by 3.9)}, \\ p_{02}(y, v) &= (a-w-v+y)E_n(w, a-w, v+1, y) + (w-y)E_n(w, a-w, v+1, y+1) \\ &= (a-v-n)E_n(w, a-w, v, y) \qquad \text{(by 3.10)}, \end{aligned}$$

and similarly

$$p_{30}(x, v) = (v-m)E_m(w, b-w, b-v, x) \qquad \text{(by 3.10)},$$

$$p_{03}(y, v) = (v-n)E_n(w, a-w, v, y) \qquad \text{(by 3.9)}.$$

Thus we have separated the variables, and obtained the equation $T(fg_m^b g_n^a) = (T_1 f)(g_m^b g_n^a)$, where

$$\begin{aligned} T_1 f(v) = {} & [v(a+b-2v) - m(b+1-m) - n(a+1-n)]f(v) \\ & + (v-n)(v-m)f(v-1) + (v-a+n)(v-b+m)f(v+1). \end{aligned}$$

By Theorem 2.3, the eigenfunctions of T_1 with eigenvalues $[ab-k(a+b+1-k)]$, $k = m+n, m+n+1, \cdots$, yield the desired functions (as $fg_m^b g_n^a$).

Set $u = v-n$, and compare the expression for $T_1 f$ to the difference equation eigenvalue problem (3.12). Observe that the solutions are of the form $E_l(a', b', c', u)$ where $b'-c' = n-m$, and either $a' = a-2n$, $c' = b-m-n$, $b' = b-2m$ or $a' = b-m-n$, $c' = a-2n$, $b' = a-m-n$ ($l = 0, 1, 2, \cdots$). These two choices determine the same function because of the symmetries of E_l (3.4), and indeed the substitution $u = v-m$ would do this also (by 3.5). We choose $a' = a-2n$, $b' = b-2m$, $c' = b-m-n$ because it reduces to the right Hahn polynomial (see Corollary 2.2 and (3.6(i))) for $m = n = 0$. Set $f(v) = E_l(a-2n, b-2m, b-m-n, v-n)$ and use (3.11) to obtain $T_1 f = [ab-(l+m+n)(a+b+1-l-m-n)]f$, thus $fg_m^b g_n^a \in P_{mn}V_{m+n+l}$, by Theorem 2.3.

For what values of l are nonzero solutions possible? The discussion after (3.12) shows that $0 \leqq l \leqq (a-2n) \wedge (b-2m) \wedge (b-m-n) \wedge (a-m-n)$ for nontrivial solutions. Setting $l = k-m-n$, we see these are the constraints specified by the representation theory (4.1), namely $l \geqq 0$ is $m+n \leqq k$, $l \leqq a-2n$ is $k-a \leqq m-n$, and $l \leqq b-2m$ is $m-n \leqq b-k$ (the constraints $2m \leqq b$, $2n \leqq a$ were specified in $g_m^b g_n^a$). We summarize these results:

THEOREM 4.5. *For integers m, n, k, w such that $0 \leqq m \leqq b/2$, $0 \leqq n \leqq a/2$, $k-a \leqq m-n \leqq b-k$, $m+n \leqq k \leqq b$, $0 \leqq w \leqq b$, the K_w-invariant functions in $P_{mn}V_k$ are spanned by $\phi_{mnk}(\cdot\,; w)$, where $\phi_{mnk}(x, y, v; w) = E_{k-m-n}(a-2n, b-2m, b-m-n, v-n)E_m(w, b-w, b-v, x)E_n(w, a-w, v, y)$, $x(\zeta) = |\zeta \cap [1, w]|$, $y(\zeta) = |\zeta \cap [a+b-w+1, a+b]|$, $v(\zeta) = |\zeta \cap [b+1, a+b]|$, $\zeta \in X$.*

We will need the L^2-norm of ϕ_{mnk} (in $L(X)$). The number of points in X having given values of x, y, v is $\binom{w}{x}\binom{b-w}{b-v-x}\binom{a-w}{v-y}\binom{w}{y}$, and this is to be divided by $\binom{a+b}{b}$ to get the normalized measure at (x, y, v).

PROPOSITION 4.6. *For integers m, n, k, w as in Theorem 4.5, $\|\phi_{mnk}(\cdot\,; w)\|^2 = A/B$, where $A = (-w)_m(w-b)_m(-w)_n(w-a)_n m!n!(k-m-n)!(-a)_k(-b)_k(a+b-m-n-k)!(-1)^{m+n}(n-a)_{k-m}(m-b)_{k-n}(a+b-k-m-n+1)(a-n+1)(b-m+1)$, and $B = (a+b)!(a+b-2k+1)(a-2n+1)(b-2m+1)$.*

Proof. This is a routine tedious calculation. Sum first over x and y using (3.7), obtaining the sum in v as $\sum_{v=m\vee n}^{(a-n)\wedge(b-m)} \binom{b-2m}{b-m-v}\binom{a-2n}{v-n} E_{k-m-n}(a-2n,\ b-2m,\ b-m-n,\ v-n)^2$ which is also done by (3.7). □

5. The addition theorem. We first give the abstract setting for the trick used to get the addition theorem. Suppose temporarily that G is an arbitrary compact group, with closed subgroup H, and further that T is a continuous unitary irreducible representation of G on an N-dimensional vector space of functions on $H\backslash G$, such that $T|H$ contains 1_H exactly once. Let $\omega=\{H\}$ be the base point in the homogeneous space $H\backslash G$.

Let A be the set of equivalence classes of unitary irreducible representations of H appearing in $T|H$, with $0\in A$ corresponding to 1_H. Thus we can write $V=\sum_{\alpha\in A}\oplus V_\alpha$, where the representation α of H is realized on the subspace V_α (possibly with multiplicity >1). Let P_α be the orthogonal projection onto V_α, thus P_α commutes with $T(h)$, $(h\in H)$. Let ϕ be the spherical function in V, that is $\phi(\omega)=1$ and $\phi(\xi h)=\phi(\xi)(\xi\in H\backslash G,\ h\in H)$, thus ϕ spans V_0. The addition formula for ϕ is

$$\phi(\xi g)=\sum_{\alpha\in A}(P_\alpha(R(g)\phi))(\xi)\qquad (\xi\in H\backslash G,\ g\in G),$$

(put $\xi=\omega g_1 h$, $g_1\in G$, $h\in H$ to see this is the usual addition formula for spherical functions). Pick an orthonormal basis for V which is a union of bases for V_α, so that the matrix representation for $T|H$ diagonalizes, and in such a way that $T_{11}(h)=1$, $T_{1i}(h)=T_{i1}(h)=0$, $2\leqq i\leqq N$ $(h\in H)$. Then $\phi(\omega g)=T_{11}(g)(g\in G)$ and $\{T_{1j}\}_{j=1}^N$ is an orthogonal basis for V, since $T_{1j}(hg)=\sum_i T_{1i}(h)T_{ij}(g)=T_{1j}(g)(h\in H,\ g\in G)$, and thus T_{1j} is a function on $H\backslash G$.

Fix $\alpha\in A$, and let J be the set of indices of the basis elements in V_α. Any $f\in V$ can be expressed in the form $\sum_{j=1}^N c_jT_{1j}$, and then $P_\alpha f=\sum_{j\in J}c_jT_{1j}$. Now apply this to $\mathrm{R}(g)\phi$. Define $f_\alpha(g_1,g_2)=(P_\alpha(\mathrm{R}(g_1)\phi))(\omega g_2)(g_1,g_2\in G)$; then:

(5.1) (i) $f_\alpha(g_1,g_2)=\sum_{j\in J}T_{j1}(g_1)T_{1j}(g_2)=\sum_{j\in J}\overline{T_{1j}(g_1^{-1})}T_{1j}(g_2)$,

since $\mathrm{R}(g_1)\phi(\omega g_2)=T_{11}(g_2g_1)=\sum_{j=1}^N T_{1j}(g_2)T_{j1}(g_1)$, and T is unitary;

(5.1)

(ii) $\int_G|f_\alpha(g_1,g_2)|^2\,dg_2=\sum_{j\in J}|T_{1j}(g_1^{-1})|^2/N$
$=f_\alpha(g_1,g_1^{-1})/N$ (Peter–Weyl theorem);

(iii) $f_\alpha(g_1h_1,h_2g_2)=f_\alpha(g_1,g_2)(h_1,h_2\in H)$, by (i);

(iv) $f_\alpha(g_1,g_2h)=f_\alpha(hg_1,g_2)(h\in H)$ since $\mathrm{R}(h)$ commutes with P_α;

(note (iii) and (iv) show it suffices to consider values of g_1,g_2 from sets of representatives of $H\backslash G/H$ and $H\backslash G$ respectively).

(5.1) (v) $f_\alpha(g_1,g_2k)=f_\alpha(g_1,g_2)$ if $k\in(g_1Hg_1^{-1})\cap H$,

indeed let $k=g_1hg_1^{-1}$, $k,h\in H$; then $f_\alpha(g_1,g_2k)=f_\alpha(kg_1,g_2)=f_\alpha(g_1h,g_2)=f_\alpha(g_1,g_2)$.

Suppose that for each $g_1\in G$, there is a unique (up to scalar multiplication) element in V_α which is invariant under $g_1Hg_1^{-1}\cap H$, that is, there is a function $F(g_1,\xi)(\xi\in H\backslash G)$ such that $\xi\mapsto F(g_1,\xi)$ is in V_α, $F(g_1,\xi k)=F(g_1,\xi)$ $(k\in g_1Hg_1^{-1}\cap H)$ and any other function having these properties is a multiple (independent of ξ) of F. This, together with (5.1(v)) shows there is a number depending on g_1, say $F_1(g_1)$, such that $f_\alpha(g_1, g_2) = F_1(g_1)F(g_1, \omega g_2)$. The following equation shows that F determines F_1:

$$f_\alpha(g_1, g_2) = \overline{F(g_1, \omega g_1^{-1})}F(g_1, \omega g_2) \Big/ \left(N \int_{H\backslash G} |F(g_1, \xi)|^2 \, d\xi\right). \tag{5.2}$$

To prove this, we apply (5.1(ii)) to F_1F, obtaining $|F_1(g_1)|^2 \int_{H\backslash G} |F(g_1, \xi)|^2 \, d\xi = F_1(g_1)F(g_1, \omega g_1^{-1})/N$, and solve for $F_1(g_1)$ (note that $\int_{H\backslash G} f(\xi)\, d\xi = \int_G f(\omega g)\, dg$ for any continuous function on $H\backslash G$).

We return to the original groups G, H as in § 1.1, to apply the above method to find the addition formula. As representatives for $H\backslash G/H$ we will use the involutions π_w, $w = 0, 1, \cdots, b$, where π_w is the product of transpositions $(1, a+b)(2, a+b-1)\cdots(w, a+b+1-w)$ for $w > 0$, $\pi_0 = 1$. The invariance group from (5.1(v)), $\pi_w H \pi_w \cap H$ is the group K_w (see § 1.1). In Theorem 4.5 we determined the unique K_w-invariant function in $P_{mn}V_k$, namely $\phi_{mnk}(x, y, v; w)$ (for the usual values of m, n, k).

The coordinates (x, y, v) for $\omega\pi_w$ are $x = 0$, $y = v = w$, so in the notation of (5.1) $F(\pi_w, \omega\pi_w) = E_{k-m-n}(a-2n,\ b-2m,\ b-m-n,\ w-n)E_m(w,\ b-w,\ b-w,\ 0)E_n(w,\ a-w,\ w,\ w) = (-1)^m(-w)_m(w-b)_m(w-a)_n(-w)_nE_{k-m-n}(a-2n,\ b-2m,\ b-m-n,\ w-n)$ (by (3.8)).

LEMMA 5.1. *For integers m, n, k, w such that $0 \leqq m \leqq b/2$, $0 \leqq n \leqq a/2$, $k-a \leqq m-n \leqq b-k$, $m+n \leqq k \leqq b$, $0 \leqq w \leqq b$,*

$$\begin{aligned}(P_{mn}\mathrm{R}(\pi_w)\phi_k)(\zeta) = {} & \frac{(-1)^n(-k)_{m+n}(k-a-b-1)_{m+n}(b-2m+1)(a-2n+1)}{n!m!(-a)_k(-b)_k(n-a)_{k-m}(m-b)_{k-n}(b-m+1)(a-n+1)} \\ & \cdot E_{k-m-n}(a-2n, b-2m, b-m-n, v-n) \\ & \cdot E_{k-m-n}(a-2n, b-2m, b-m-n, w-n) \\ & \cdot E_m(w, b-w, b-v, x)E_n(w, a-w, v, y),\end{aligned}$$

where $x = |\zeta \cap [1, w]|$, $y = |\zeta \cap [a+b+1-w, a+b]|$, $v = |\zeta \cap [b+1, a+b]|$, $\zeta \in X$.

Proof. This is proved by (5.2) and the L^2-norm from Proposition 4.6. Recall that the dimension of V_k is $\binom{a+b}{k}\left(\frac{a+b-2k+1}{a+b-k+1}\right)$. The cancellation of the factors like $(-w)_m$, $(w-b)_m$ etc., is permissible since E_mE_n is zero unless $m \leqq w \wedge (b-w)$, $n \leqq w \wedge (a-w)$. □

The addition formula for ϕ_k is $\phi_k(\zeta\pi_w) = \sum_{m,n} (P_{mn}\mathrm{R}(\pi_w)\phi_k)(\zeta)$. The left hand side is a Hahn polynomial evaluated at $|\zeta\pi_w \cap [b+1, a+b]| = |\zeta \cap [b+1, a+b]\pi_w| = |\zeta \cap ([b+1, a+b-w] \cup [1, w])| = v - y + x$. To get a more symmetric expression, replace x by $w-x$ and use the identity $E_m(w, b-w, b-v, w-x) = (-1)^m E_m(w, b-w, v, x)$ (3.5)). To obtain the addition theorem we sum the terms of Lemma 5.1 over m, n.

THEOREM 5.2. *For integers a, b, k such that $0 \leqq k \leqq b \leqq a$,*

$$\begin{aligned}Q_k(v+w-x-y; -a-1, -b-1, b) = {} & \sum_{m,n} c_{mnk}(a, b) \\ & \cdot E_{k-m-n}(a-2n, b-2m, b-m-n, w-n) \\ & \cdot E_{k-m-n}(a-2n, b-2m, b-m-n, v-n) \\ & E_m(w, b-w, v, x)E_n(w, a-w, v, y),\end{aligned} \tag{5.3}$$

where

$$c_{mnk}(a, b)=\frac{(-1)^{m+n}(-k)_{m+n}(k-a-b-1)_{m+n}(b-2m+1)(a-2n+1)}{n!m!(-a)_k(-b)_k(n-a)_{k-m}(m-b)_{k-n}(b-m+1)(a-n+1)},$$

and the sum is taken over integers m, n with $0 \leqq m+n \leqq k$, $0 \leqq m \leqq b/2$, $0 \leqq n \leqq a/2$, $k-a \leqq m-n \leqq b-k$, and v, w, x, y are integers such that $v, w \in [0, b]$, $0 \vee (w+v-b) \leqq x \leqq v \wedge w$ and $0 \vee (w+v-a) \leqq y \leqq v \wedge w$.

COROLLARY 5.3. *Suppose b, k are integers, $0 \leqq k \leqq b$ and a is real, $a \neq 0$, $1, \cdots 2b-1$; then the formula* (5.3) *holds, where the sum is taken over $0 \leqq m+n \leqq k$, $0 \leqq m \leqq b/2$, $n \geqq 0$, $m-n \leqq b-k$ and v, w, x, y are integers such that $w, v \in [0, b]$, $0 \vee (w+v-b) \leqq x \leqq w \wedge v$ and $0 \leqq y \leqq w \wedge v$.*

Proof. Applying the theorem to integer values of $a \geqq 2b$, we obtain the constraints as stated; note $n \leqq k \leqq b \leqq a/2$, $m-n \geqq -k \geqq k-2b \geqq k-a$, and $w+v-a \leqq 2b-a \leqq 0$. The right side of (5.3) is a meromorphic function in a, with poles at most at $0, 1, \cdots, 2b-1$. □

By using the orthogonality relations of the E-functions we can obtain the product formula for Q_k, annihilating all but the $m=n=0$ term in (5.3). Indeed let b, k be integers with $0 \leqq k \leqq b$; then

$$\begin{aligned} &Q_k(v; -a-1, -b-1, b)Q_k(w; -a-1, -b-1, b) \\ &= \sum_x \sum_y K(x, y; v, w; a, b)Q_k(v+w-x-y; -a-1, -b-1, b) \end{aligned} \tag{5.4}$$

where

$$K(x, y; v, w; a, b)=\frac{(-w)_x(-w)_y(w-b)_{v-x}(w-a)_{v-y}v!v!}{x!(v-x)!y!(v-y)!(-b)_v(-a)_v},$$

x, y are integers with $0 \vee (w+v-b) \leqq x \leqq w \wedge v$, and $0 \vee (w+v-a) \leqq y \leqq w \wedge v$ for $a=b, b+1, \cdots, 2b-1$ but $0 \leqq y \leqq w \wedge v$ for real $a \neq 0, 1, \cdots, 2b-1$. Observe $K \geqq 0$ for $a=b, \cdots, 2b-1$ and real $a>2b-1$.

For the product formula for $Q_k(\cdot; -a-1, -b-1, c)$ with a, b, c integers with $c \leqq a \wedge b$, see Dunkl [5] (and for a proof not using groups see Rahman [8]). If $c=b$ these are spherical functions, but if $b>c$ then they are intertwining functions, certain functions on the double coset space $(S_c \times S_{a+b-c}) \backslash S_{a+b} / (S_b \times S_a)$.

REFERENCES

[1] W. BAILEY, *Generalized Hypergeometric Series*, Cambridge University Press, New York, 1935.

[2] P. DELSARTE, *An algebraic approach to the association schemes of coding theory*, Philips Research Reps. Suppl., 10, 1973.

[3] ———, *Hahn polynomials, discrete harmonics and t-designs*, Rep. R295, MBLE, Brussels, April 1975.

[4] C. DUNKL, *A Krawtchouk polynomial addition theorem and wreath products of symmetric groups*, Indiana Univ. Math. J., 25 (1976), pp. 335–358.

[5] ———, *Spherical functions on compact groups and applications to special functions*, Symposia Mathematica, Vol. 22, Rome, 1977.

[6] G. GASPER, *Projection formulas for orthogonal polynomials of a discrete variable*, J. Math. Anal. Appl., 45 (1974), pp. 176–198.

[7] S. KARLIN AND J. MCGREGOR, *The Hahn polynomials, formulas and an application*, Scripta Math., 26 (1961), pp. 33–46.

[8] M. RAHMAN, *A positive kernel for Hahn-Eberlein polynomials*, to appear.

[9] G. DE B. ROBINSON, *Representation Theory of the Symmetric Group*, University of Toronto Press, Toronto, 1961.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-141C/78/0904-0007 $01.00/0

MINIMUM PRINCIPLES FOR ILL-POSED PROBLEMS*

JOEL N. FRANKLIN†

Abstract. Ill-posed problems $Ax = h$ are discussed in which A is Hermitian and postive definite; a bound $\|Bx\| \leqq \beta$ is prescribed. A minimum principle is given for an approximate solution $\hat{x}$. Comparisons are made with the least-squares solutions of K. Miller, A. Tikhonov, et al. Applications are made to deconvolution, the backward heat equation, and the inversion of ill-conditioned matrices. If A and B are positive-definite, commuting matrices, the approximation $\hat{x}$ is shown to be about as accurate as the least-squares solution and to be more quickly and accurately computable.

1. Introduction. This paper discusses ill-posed problems of the form

$$Ax = h, \tag{1.1}$$

where A is a positive-definite Hermitian operator mapping a Hilbert space H into itself. Although we assume $(Ax, x) > 0$ if $x \neq 0$, we often assume also that $\|Au\|$ may be arbitrarily near zero on the unit sphere $\|u\| = 1$. Then A cannot have a bounded inverse, and the problem (1.1) is ill-posed because the solution x, if it exists, is unstable: arbitrarily small perturbations of the data, h, can produce arbitrarily large perturbations of the solution, x. Typical of such problems is the Fredholm integral equation of the first kind:

$$\int_0^1 A(s, t)x(t)\, dt = h(s) \qquad (0 < s < 1), \tag{1.2}$$

where $A(s, t)$ is bounded, integrable, self-adjoint, and positive definite.

We shall also consider equations of the form (1.1) where A is an $n \times n$ positive-definite Hermitian matrix, and where the data h and the solution x lie in the n-dimensional vector space. In practice, this problem is ill-posed if A has a large condition number, which is defined as the ratio of largest to smallest eigenvalues. Here a bounded inverse A^{-1} does exist in theory, but the solution $x = A^{-1}h$ is numerically unstable because the relative error

$$\frac{\|\delta x\|}{\|x\|} \div \frac{\|\delta h\|}{\|h\|} \tag{1.3}$$

may become large. In fact, the maximum value of the relative error equals the condition number.

Let x^0 be the unknown solution, and let h be numerical or other approximate data satisfying

$$\|Ax^0 - h\| \leqq \varepsilon, \tag{1.4}$$

where ε is small but positive. Here we have replaced the equation $Ax = h$ by an inequality, which is more realistic because it admits the possibility of a nonzero data error. As originally shown by C. Pucci [16], such a problem can often be regularized by additional information in the form of a prescribed bound

$$\|Bx^0\| \leqq \beta. \tag{1.5}$$

Here the operator B and the finite bound β are known. This is new, given information, which is independent of the original information (1.4).

*Received by the editors October 15, 1976, and in revised form July 28, 1977.
† Firestone Laboratory, California Institute of Technology, Pasadena, California 91125.

Keith Miller [14] has considered the problem (1.4), (1.5), in which the linear operators A and B are not required to be Hermitian, but are required to be bounded. For such problems he has given several very useful numerical methods based on the least-squares principle

$$\|Ax-h\|^2+\lambda^2\|Bx\|^2=\text{minimum}. \tag{1.6}$$

If ε and β are known explicitly, the preferred choice of λ is $\lambda=\varepsilon/\beta$. The minimal solution is

$$x^1=(A^*A+\lambda^2B^*B)^{-1}A^*h, \tag{1.7}$$

which is the solution by Miller's Method 1.

For the problem of inverting ill-conditioned matrices similar formulas, making use of a prescribed bound, have been used since 1959 or earlier; see references [4] through [8], [12], [13], [15], and the book by C. Lawson and R. Hanson [10, pp. 188–194].

For the ill-posed Fredholm equation (1.2) A. N. Tikhonov [17] developed a least-squares method. Here the prescribed bound takes the form

$$\Omega^2(x)\equiv\int_0^1[x^2(t)+\dot{x}^2(t)]\,dt\leqq\beta^2. \tag{1.8}$$

Or one may use any other Sobolev norm for $\Omega(x)$. Tikhonov's minimum principle for an approximate solution is this:

$$\|Ax-h\|^2+\lambda^2\Omega^2(x)=\text{minimum}. \tag{1.9}$$

Error analysis in general and for certain applications has been given in [3].

Tikhonov's minimum principle (1.9) can be put in Miller's form (1.6) if B is suitably defined, but now B is unbounded. For example, we may define B on the domain of functions

$$x(t)=\sum_{n=0}^{\infty}a_n\cos n\pi t \qquad (0<t<1), \tag{1.10}$$

where $\sum n^2a_n^2<\infty$. Then we define

$$Bx(t)=\sum_{n=0}^{\infty}(1+n^2\pi^2)^{1/2}a_n\cos n\pi t. \tag{1.11}$$

This makes B positive definite and unbounded, with domain dense in the real Hilbert space L^2; and

$$\int_0^1[x^2(t)+\dot{x}^2(t)]\,dt=\|Bx\|^2=a_0^2+\frac{1}{2}\sum_1^{\infty}(1+n^2\pi^2)a_n^2. \tag{1.12}$$

Now the Tikhonov principle (1.9) takes Miller's form (1.6), and Tikhonov's minimal solution is given by (1.7).

In the present paper, we will analyze a different minimum principle for the ill-posed problem (1.4) with prescribed bound (1.5). Though A is bounded, we allow B to be unbounded (as it must be to include Tikhonov's regularizations); but we require B^{-1} to be bounded. We define $\hat{x}$ to be the solution of this problem:

$$(Ax,x)-2\,\mathrm{Re}\,(h,x)+\lambda(Bx,x)=\text{minimum}, \tag{1.13}$$

where $\lambda = \varepsilon/\beta$. The solution has the simple form

$$\hat{x} = (A + \lambda B)^{-1} h. \tag{1.14}$$

This principle is less generally applicable than Miller's, since it applies only to ill-posed problems $Ax = h$ in which A is Hermitian and positive definite. But the simple form of the solution $\hat{x}$ has advantages in numerical analysis, particularly in the inversion of ill-conditioned matrices.

For matrices, both principles are examples of least squares; see Lawson and Hanson [10]. The principle (1.6) comes from the least-squares problem

$$\begin{pmatrix} A \\ \lambda B \end{pmatrix} x = \begin{pmatrix} h \\ 0 \end{pmatrix}. \tag{1.15}$$

The principle (1.13) comes from the least-squares problem

$$\begin{pmatrix} L \\ \sqrt{\lambda} R \end{pmatrix} x = \begin{pmatrix} (L^*)^{-1} h \\ 0 \end{pmatrix} \tag{1.16}$$

where L and R appear in the Cholesky factorizations $L^*L = A$, $R^*R = B$.

2. Error estimates. Let $\langle x \rangle$ be a seminorm on the Hilbert space H. If x^0 is the unknown solution of the inequalities

$$\|Ax^0 - h\| \leqq \varepsilon, \qquad \|Bx^0\| \leqq \beta, \tag{2.1}$$

and if x is an approximate solution, then $\langle x - x^0 \rangle$ is a measure of the error. Miller [14] defines these quantities:

$$\mathcal{M}(\varepsilon, \beta) = \sup \{\langle x \rangle : \|Ax\| \leqq \varepsilon, \|Bx\| \leqq \beta\}, \tag{2.2}$$

$$\mathcal{M}_1(\varepsilon, \beta) = \sup \{\langle x \rangle : \|Ax\|^2 + \lambda^2 \|Bx\|^2 \leqq 2\varepsilon^2\} \tag{2.3}$$

where $\lambda = \varepsilon/\beta$. In his Lemma 3, he proves

$$\mathcal{M}(\varepsilon, \beta) \leqq \mathcal{M}_1(\varepsilon, \beta) \leqq \sqrt{2}\mathcal{M}(\varepsilon, \beta) \tag{2.4}$$

(I have changed his notation by using β instead of E.)

The quantity $\mathcal{M}$ shows how much the information $\|Bx\| \leqq \beta$ restricts $\langle x \rangle$ if you know $\|Ax\| \leqq \varepsilon$. This is important because in an ill-posed problem $Ax = h$, the norm $\|Au\|$ may tend to zero on the unit sphere, $\|u\| = 1$; therefore, $\|x\|$—and perhaps $\langle x \rangle$—may be very large even if $\|Ax\| \leqq \varepsilon$.

Miller presents four numerical methods based on least squares. If both ε and β are known explicitly (and are not just known to exist), the preferred method is Method 1; and this is the method we shall use for purposes of comparison. Miller's minimum principle and its solution, x^1, appear in our formulas (1.6), (1.7). In his Lemma 4, he gives this error estimate:

$$\langle x^1 - x^0 \rangle \leqq \mathcal{M}_1(\varepsilon, \beta). \tag{2.5}$$

Our purpose is to examine the minimum principle (1.13) and the solution, $\hat{x}$, given in (1.14). *We assume that A is bounded, Hermitian, and positive definite.* We are concerned with ill-posed problems, in which A^{-1} is unbounded or very large in norm. *We assume B is Hermitian and positive definite, with a bounded inverse* B^{-1}; we assume that the domain of B is dense in the Hilbert space, but we do not assume B is bounded.

If $\langle x \rangle$ is any seminorm, we define these quantities:

$$\mathcal{N}(\varepsilon, \beta) = \sup \{\langle x \rangle : (Ax, x) \leqq \varepsilon \|x\|, (Bx, x) \leqq \beta \|x\|\}, \tag{2.6}$$

$$\mathcal{N}_1(\varepsilon, \beta) = \sup \{\langle x \rangle : (Ax, x) + \lambda (Bx, x) \leqq 2\varepsilon \|x\|\}, \tag{2.7}$$

where $\varepsilon > 0, \beta > 0$, and $\lambda = \varepsilon/\beta$. These quantities are practically the same, namely,

$$\mathcal{N}(\varepsilon, \beta) \leqq \mathcal{N}_1(\varepsilon, \beta) \leqq 2\mathcal{N}(\varepsilon, \beta). \tag{2.8}$$

The reason for defining both of them is that sometimes one is easier to compute than the other. The last three formulas are comparable to Miller's formulas that we have numbered (2.2), (2.3), and (2.4); we will obtain quantitative comparisons later. First we will estimate the error $\langle \hat{x} - x^0 \rangle$.

THEOREM 1. *Let A and B satisfy the preceding assumptions. Let x^0 satisfy* (2.1), *and let $\hat{x} = (A + \lambda B)^{-1} h$. Then $\hat{x}$ uniquely solves the minimum problem* (1.13), *and*

$$\langle \hat{x} - x^0 \rangle \leqq \mathcal{N}_1(\varepsilon, \beta). \tag{2.9}$$

Proof. The operator $(A + \lambda B)$ has a bounded inverse because

$$((A + \lambda B)x, x) \geqq \lambda (BX, x) \geqq \lambda \|B^{-1}\|^{-1} \|x\|^2. \tag{2.10}$$

Then, since $\hat{x} \equiv (A + \lambda B)^{-1} h$,

$$\begin{aligned} &(Ax, x) - 2\mathrm{Re}\,(h, x) + \lambda (Bx, x) \\ &\qquad = ((A + \lambda B)(x - \hat{x}), (x - \hat{x})) - ((A + \lambda B)\hat{x}, \hat{x}) \\ &\qquad \geqq -((A + \lambda B)\hat{x}, \hat{x}) = -(h, \hat{x}) \end{aligned} \tag{2.11}$$

with equality if and only if $x = \hat{x}$. This proves that $\hat{x}$ is the unique solution of the minimum problem (1.13).

Let $\varphi = x - \hat{x}$. Then

$$\begin{aligned} (A\varphi, \varphi) + \lambda (B\varphi, \varphi) &= ((A + \lambda B)(x - \hat{x}), \varphi) \\ &= ((A + \lambda B)x - h, \varphi). \end{aligned}$$

Thus, for all x we have the identity

$$(A\varphi, \varphi) + \lambda (B\varphi, \varphi) = (Ax - h, \varphi) + \lambda (Bx, \varphi). \tag{2.12}$$

Set $x = x^0$. Then $\|Ax - h\| \leqq \varepsilon$ and $\lambda \|Bx\| \leqq \varepsilon$, and so

$$(A\varphi, \varphi) + \lambda (B\varphi, \varphi) \leqq 2\varepsilon \|\varphi\| \qquad (\varphi = x^0 - \hat{x}). \tag{2.13}$$

This gives the error estimate (2.9). □

3. Comparisons. Now we will compare the minimum principles (1.6) and (1.13).

For all $\lambda \geqq 0$ the expression (1.6) is $\geqq 0$ and therefore has the finite lower bound 0. This is not always true of the expression (1.13). If $\lambda = 0$, it becomes

$$(Ax, x) - 2\mathrm{Re}\,(h, x). \tag{3.1}$$

If h lies outside the range of A, this expression may tend to $-\infty$ as x varies. But of course this cannot happen if $\lambda > 0$ in (1.13), since we have assumed $\|B^{-1}\| < \infty$.

As an example of (3.1), let H be the Hilbert space of vectors x with real components satisfying

$$\|x\|^2 = \sum_{n=1}^{\infty} x_n^2 < \infty.$$

Let (3.1) take the form

$$\sum_{n=1}^{\infty} n^{-1}x_n^2-2\sum_{n=1}^{\infty} n^{-1}x_n. \tag{3.2}$$

If we set $x_n=1$ for $n=1,\cdots,N$ and set $x_n=0$ for $n>N$, the expression (3.2) equals

$$-\sum_{n=1}^{N} n^{-1}\to-\infty \quad \text{as } N\to\infty. \tag{3.3}$$

Now let us compare the quantities $\mathcal{M}_1(\varepsilon,\beta)$ and $\mathcal{N}_1(\varepsilon,\beta)$, which are the upper bounds for the errors $\langle x^1-x^0\rangle$.

THEOREM 2. *Let $\mathcal{M}_1$ and $\mathcal{N}_1$ be defined as in* (2.3) *and* (2.7). *Then for every seminorm $\langle x\rangle$,*

$$\mathcal{M}_1(\varepsilon,\beta)\leqq\mathcal{N}_1(\varepsilon,\beta). \tag{3.4}$$

Moreover, the ratio $\mathcal{M}_1/\mathcal{N}_1$ may be arbitrarily near zero. But if A and B commute, and if $\langle x\rangle\equiv\|x\|$, then

$$\mathcal{N}_1(\varepsilon,\beta)\leqq\sqrt{2}\mathcal{M}_1(\varepsilon,\beta). \tag{3.5}$$

In many ill-posed problems with prescribed bounds, A and B do commute. Then this theorem shows that the errors in the two methods, $\|x^1-x^0\|$ and $\|\hat{x}-x^0\|$, have practically the same upper bound.

Proof of the theorem. If

$$\|Ax\|^2+\lambda^2\|Bx\|^2\leqq 2\varepsilon^2$$

then

$$\begin{aligned}(Ax,x)+\lambda(Bx,x)&\leqq(\|Ax\|+\lambda\|Bx\|)\|x\|\\&\leqq\sqrt{2}(\|Ax\|^2+\lambda^2\|Bx\|^2)^{1/2}\|x\|\\&\leqq 2\varepsilon\|x\|.\end{aligned} \tag{3.6}$$

That proves the inequality (3.4).

Next we will show that $\mathcal{M}_1/\mathcal{N}_1$ may go to zero. For an example, we will use the real Euclidian vector space H with n dimensions. We define the diagonal matrix

$$A=B=n^{1/2}\operatorname{diag}(1,2^{-1/2},3^{-1/2},\cdots,n^{-1/2}). \tag{3.7}$$

We define the seminorm $\langle x\rangle=\|Ax\|$. Let $\lambda=1$. Then our definitions become

$$\mathcal{M}_1(\varepsilon,\beta)=\sup\{\|Ax\|:2\|Ax\|^2\leqq 2\varepsilon^2\} \tag{3.8}$$

$$\mathcal{N}_1(\varepsilon,\beta)=\sup\{\|Ax\|:2(Ax,x)\leqq 2\varepsilon\|x\|\}. \tag{3.9}$$

Then $\mathcal{M}_1=\varepsilon$, but

$$\mathcal{N}_1=\varepsilon\max\frac{\|Ax\|\,\|x\|}{(Ax,x)}. \tag{3.10}$$

Let $x_k = n^{-1/2}(k=1,\cdots,n)$. Then as $n\to\infty$

$$\|x\|^2 = 1,$$

$$\|Ax\|^2 = \sum_{k=1}^{n} k^{-1} \sim \log n, \tag{3.11}$$

$$(Ax, x) = n^{-1/2}\sum_{k=1}^{n} k^{-1/2} \to 2.$$

Therefore, for fixed $\varepsilon > 0$, $\mathcal{N}_1 \to \infty$ as $n\to\infty$; and so $\mathcal{M}_1/\mathcal{N}_1 \to 0$.

Now suppose $\langle x\rangle \equiv \|x\|$. Then

$$\begin{aligned}\mathcal{M}_1 &= \sup\{\|x\| : \|Ax\|^2 + \lambda^2\|Bx\|^2 \leqq 2\varepsilon^2\}\\ &= \sup\{\|x\| : ((A^2+\lambda^2B^2)x, x) \leqq 2\varepsilon^2\}\\ &= \sqrt{2}\,\varepsilon\,\|(A^2+\lambda^2B^2)^{-1}\|^{1/2}.\end{aligned} \tag{3.12}$$

Similarly, we find

$$\begin{aligned}\mathcal{N}_1 &= \sup\{\|x\| : ((A+\lambda B)x, x) \leqq 2\varepsilon\|x\|\}\\ &= 2\varepsilon\|(A+\lambda B)^{-1}\|.\end{aligned} \tag{3.13}$$

If A and B commute, then the bounded operators A and $(\lambda B)^{-1}$ have spectral representations

$$\begin{aligned}A &= \int \rho_v\, dE_v,\\ (\lambda B)^{-1} &= \int \sigma_v^{-1}\, dE_v\end{aligned} \tag{3.14}$$

where dE_v is a common projection operator, and where

$$0 \leqq \rho_v \leqq \|A\|, \qquad 0 \leqq \sigma_v^{-1} \leqq \lambda^{-1}\|B^{-1}\|. \tag{3.15}$$

(If A and B are matrices, then ρ_v and σ_v are eigenvalues of A and λB belonging to a common eigenvector.)

The operators $(A+\lambda B)^{-1}$ and $(A^2+\lambda^2B^2)^{-1}$ are bounded, since the constant λ is positive. They have these spectral representations:

$$\begin{aligned}(A+\lambda B)^{-1} &= \int (\rho_v+\sigma_v)^{-1}\, dE_v,\\ (A^2+\lambda^2B^2)^{-1} &= \int (\rho_v^2+\sigma_v^2)^{-1}\, dE_v.\end{aligned} \tag{3.16}$$

Then

$$\begin{aligned}\|(A+\lambda B)^{-1}\| &= \sup\,(\rho_v+\sigma_n)^{-1},\\ \|(A^2+\lambda^2B^2)^{-1}\| &= \sup\,(\rho_v^2+\sigma_v^2)^{-1}.\end{aligned} \tag{3.17}$$

But for all positive ρ and σ

$$(\rho+\sigma)^{-1} \leqq (\rho^2+\sigma^2)^{-1/2},$$

and so

$$\|(A+\lambda B)^{-1}\| \leqq \|(A^2+\lambda^2B^2)^{-1}\|^{1/2}. \tag{3.18}$$

Now (3.12) and (3.13) imply $\mathcal{N}_1 \leqq \sqrt{2}\,\mathcal{M}_1$. □

Note 1. Our proof of the inequality $\mathcal{N}_1 \leqq \sqrt{2}\,\mathcal{M}_1$ assumes that A and B commute. If A and B do not commute, the inequality may be false. For example, let $\lambda = 1$ and let

$$A = \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}, \qquad B = \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix}.$$

Then

$$A+B = \begin{pmatrix} 6 & -4 \\ -4 & 6 \end{pmatrix}, \qquad A^2+B^2 = \begin{pmatrix} 34 & -24 \\ -24 & 34 \end{pmatrix}.$$

The minimum eigenvalues of these two matrices are

$$\lambda_{\min}(A+B) = 2, \qquad \lambda_{\min}(A^2+B^2) = 10.$$

Therefore,

$$\|(A+B)^{-1}\| = 1/2 > \|(A^2+B^2)^{-1}\|^{1/2} = 1/\sqrt{10}.$$

Thus, for this example the inequality (3.18) is false, and $\mathcal{N}_1 > \sqrt{2}\,\mathcal{M}_1$.

Note 2. In Miller's assumption $\|Bx^0\| \leqq \beta$, where he takes B to be bounded, we make no loss of generality by assuming $B = B^*$, since we can always replace B by the Hermitian operator $B_1 = (B^*B)^{1/2}$ and then assume $\|B_1 x^0\| \leqq \beta$.

Note 3. If A has an inverse, the principle (1.13) can be put in the form (1.6). If we define

$$A_1 = A^{1/2}, \quad g = A^{-1/2}h, \quad \lambda_1 = \lambda^{1/2}, \quad B_1 = B^{1/2}$$

then the principle

$$(Ax, x) - 2\mathrm{Re}\,(h, x) + \lambda\,(Bx, x) = \text{minimum}$$

takes the form

$$\|A_1 x - g\|^2 + \lambda_1^2 \|B_1 x\|^2 = \text{minimum}.$$

But this form cannot be used if A lacks a bounded inverse, which is the case if the original problem $Ax = h$ is ill-posed.

4. Deconvolution. We will now apply our results to the real convolution equation

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} a(z-y)u(y)\,dy = h(z) \qquad (-\infty < z < \infty). \tag{4.1}$$

(Here we have called the unknown u instead of x.) Let the function $a(z)$ have the Fourier transform

$$A(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega z} a(z)\,dz \qquad (-\infty < \omega < \infty), \tag{4.2}$$

and let u and h have the Fourier transforms U and H. Then the convolution equation (4.1) becomes

$$A(\omega)U(\omega) = H(\omega). \tag{4.3}$$

As a rule, this equation is ill-posed. A data error $\delta H(\omega)$ produces a solution error

$$\delta U(\omega) = A^{-1}(\omega)\,\delta H(\omega). \tag{4.4}$$

If the transform $A(\omega)\to 0$ as $\omega\to\pm\infty$, then a data error at high frequencies is greatly magnified when it is multiplied by $A^{-1}(\omega)$.

We shall suppose $A(\omega)>0$. This means that the original convolution operator in (4.1) is positive definite, since

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(z)a(z-y)u(y)\,dy\,dz=\int_{-\infty}^{\infty} A(\omega)|U(\omega)|^2\,d\omega. \tag{4.5}$$

Although $A(\omega)>0$, we shall usually have $A(\omega)\to 0$ as $\omega\to\pm\infty$.

We look for an unknown solution $U_0(\omega)$. We replace the equation (4.3) by an inequality

$$\|A(\omega)U_0(\omega)-H(\omega)\|\leqq \varepsilon. \tag{4.6}$$

This permits a nonzero data error $\delta H(\omega)$ with L^2 norm $\leqq\varepsilon$.

The problem is still ill-posed; to make sense of it we need some new given information. We shall suppose this information takes the form of a prescribed bound

$$\|B(\omega)U_0(\omega)\|\leqq\beta, \tag{4.7}$$

where $B(\omega)>0$ *and* $B(\omega)$ has a positive lower bound. In fact, we shall usually have $B(\omega)\to\infty$ as $\omega\to\pm\infty$. In any case, the inverse $B^{-1}(\omega)$ is bounded.

The least-squares approach to the extended problem (4.6), (4.7) is to solve

$$\|AU-H\|^2+\lambda^2\|BU\|^2=\text{minimum}, \tag{4.8}$$

where $\lambda=\varepsilon/\beta$. The solution is

$$U_1(\omega)=\frac{A(\omega)H(\omega)}{A^2(\omega)+\lambda^2B^2(\omega)}. \tag{4.9}$$

This is the approach taken by Miller, although we have here allowed B to be unbounded. The inverse transform

$$u_1(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-iz\omega}U_1(\omega)\,d\omega \tag{4.10}$$

is an approximate solution of the original convolution equation (4.1).

A second approach to (4.6), (4.7) is to solve

$$(AU, U)-2(U, H)+\lambda(BU, U)=\text{minimum}, \tag{4.11}$$

where the inner product is defined by

$$(F, G)=\int_{-\infty}^{\infty} F(\omega)\bar{G}(\omega)\,d\omega. \tag{4.12}$$

(The expression (4.11) is real-valued, since we assume $U(\omega)$ and $H(\omega)$ are the transforms of real-valued functions $u(z)$ and $h(z)$.) The solution of (4.11) is

$$\hat{U}(\omega)=\frac{H(\omega)}{A(\omega)+\lambda B(\omega)}. \tag{4.13}$$

The inverse transform $\hat{u}(z)$ is a second approximate solution of (4.1).

A third approach is to use what Miller calls the method of *partial eigenfunction expansion*. This is not a minimum principle, but it is useful for comparison; and it is often a good numerical method. Define the set

$$\Omega = \{\omega : A(\omega) \geqq \lambda B(\omega)\}, \tag{4.14}$$

and let Ω' be its complement. Then define the *cutoff* solution

$$U_c(\omega) = \begin{cases} A^{-1}(\omega)H(\omega) & \text{on } \Omega, \\ 0 & \text{on } \Omega'. \end{cases} \tag{4.15}$$

We now have three approximate solutions $U(\omega)$, and we can compare their errors $\|U - U_0\|$, where U_0 is the unknown true solution of (4.6), (4.7). In the present application, the general formulas (3.12) and (3.13) imply

$$\mathcal{M}_1(\varepsilon, \beta) = \frac{\sqrt{2}\varepsilon}{\inf_\omega [A^2(\omega) + \lambda^2 B^2(\omega)]^{1/2}}, \tag{4.16}$$

$$\mathcal{N}_1(\varepsilon, \beta) = \frac{2\varepsilon}{\inf_\omega [A(\omega) + \lambda B(\omega)]}, \tag{4.17}$$

and formulas (3.4) and (3.5) state

$$\mathcal{M}_1 \leqq \mathcal{N}_1 \leqq \sqrt{2}\mathcal{M}_1. \tag{4.18}$$

For the approximation $U_1(\omega)$ Miller's error bound is

$$\|U_1 - U_0\| \leqq \mathcal{M}_1(\varepsilon, \beta). \tag{4.19}$$

For the approximation $\hat{U}(\omega)$ our error bound is

$$\|U_2 - U_0\| \leqq \mathcal{N}_1(\varepsilon, \beta). \tag{4.20}$$

For the cutoff approximation, $U_c(\omega)$, Miller's error bound is

$$\|U_c - U_0\| \leqq \sqrt{2}\mathcal{M}(\varepsilon, \beta) \leqq \sqrt{2}\mathcal{M}_1(\varepsilon, \beta), \tag{4.21}$$

where

$$\mathcal{M}(\varepsilon, \beta) = \sup \{\|F\| : \|AF\| \leqq \varepsilon, \|BF\| \leqq \beta\}.$$

The last estimate appears in Miller's Lemma 8 in [14]. He proves it by the theory of spectral representation for commuting bounded operators. Since in our application B is usually unbounded, we should give a separate proof. We will prove

$$\|A(U_c - U_0)\| \leqq \sqrt{2}\varepsilon, \tag{4.22}$$

$$\|B(U_c - U_0)\| \leqq \sqrt{2}\beta. \tag{4.23}$$

These inequalities directly imply (4.21).

For us, the operators A and B are just ordinary positive functions, and the proofs are easy. For any $F(\omega)$, we have

$$\begin{aligned} \|F\|^2 &\equiv \int_{-\infty}^{\infty} |F|^2 \, d\omega = \int_\Omega |F|^2 \, d\omega + \int_{\Omega'} |F|^2 \, d\omega \\ &\equiv \|F\|_\Omega^2 + \|F\|_{\Omega'}^2. \end{aligned} \tag{4.24}$$

With this notation, we find

$$\|A(U_c-U_0)\|^2=\|A(U_c-U_0)\|_\Omega^2+\|A(U_c-U_0)\|_{\Omega'}^2 \tag{4.25}$$
$$=\|H-AU_0\|_\Omega^2+\|AU_0\|_{\Omega'}^2.$$

Since $A<\lambda B$ on Ω', we find

$$\|AU_0\|_{\Omega'}\leqq\lambda\|BU_0\|_{\Omega'}\leqq\lambda\beta=\varepsilon. \tag{4.26}$$

Now (4.26) yields (4.22). Similarly,

$$\|B(U_c-U_0)\|^2=\|B(U_c-U_0)\|_\Omega^2+\|B(U_c-U_0)\|_{\Omega'}^2 \tag{4.27}$$
$$\leqq\lambda^{-2}\|A(U_c-U_0)\|_\Omega^2+\|BU_0\|^2\leqq 2\beta^2.$$

This proves (4.23), and now (4.21) follows.

All three approximations have the same form, namely,

$$U(\omega)=(1+[\lambda B(\omega)A^{-1}(\omega)]^p)^{-1}A^{-1}(\omega)H(\omega). \tag{4.28}$$

We get $U_1(\omega)$, $\hat{U}(\omega)$, $U_c(\omega)$ for $p=2, 1, \infty$.

Example. We will consider the backward heat equation with prescribed bound at a previous time. Let the temperature $\varphi(z, t)$ solve

$$\frac{\partial\varphi}{\partial t}=\frac{\partial^2\varphi}{\partial z^2}\qquad(-\infty<z<\infty). \tag{4.29}$$

If $\tau>0$ is fixed, and if $\varphi(z,\tau)$ is given, we wish to compute the initial temperature $\varphi(z, 0)$. This problem is ill-posed.

If we set $h(z)=\varphi(z,\tau)$ and $u(z)=\varphi(z,0)$, we can state this problem as a convolution equation (4.1) by defining the kernel

$$a(z)=\frac{1}{\sqrt{2\tau}}e^{-z^2/(4\tau)}\qquad(-\infty<z<\infty). \tag{4.30}$$

Or we can use the equation (4.3) for the transforms:

$$e^{-\tau\omega^2}U(\omega)=H(\omega). \tag{4.31}$$

More realistically, we look for a solution $U_0(z)$ to an inequality:

$$\|e^{-\tau\omega^2}U_0(\omega)-H(\omega)\|\leqq\varepsilon. \tag{4.32}$$

In the extended problem we have additional information: at a previous time, $t=-\sigma<0$, we have

$$\|\varphi(z,-\sigma)\|\leqq\beta. \tag{4.33}$$

In terms of Fourier transforms, this says

$$\|e^{\sigma\omega^2}U(\omega)\|\leqq\beta. \tag{4.34}$$

In this example,

$$A(\omega)=e^{-\tau\omega^2},\qquad B(\psi)=e^{\sigma\omega^2}. \tag{4.35}$$

The three approximations are (with $\lambda=\varepsilon/\beta$)

$$U_1(\omega)=\frac{e^{-\tau\omega^2}H(\omega)}{e^{-2\tau\omega^2}+\lambda^2e^{2\sigma\omega^2}}, \tag{4.36}$$

(4.37) $$U(\omega)=\frac{H(\omega)}{e^{-\tau\omega^2}+\lambda\, e^{\sigma\omega^2}},$$

and

(4.38) $$U_c(\omega)=\begin{cases} e^{\tau\omega^2}H(\omega) & \text{if } |\omega|\leqq\omega_c,\\ 0 & \text{if } |\omega|>\omega_c,\end{cases}$$

where the cutoff frequency ω_c is the positive root of the equation

(4.39) $$e^{-\tau\omega^2}=\lambda\, e^{\sigma\omega^2}.$$

As we have seen, all the error estimates are about the same; they appear in formulas (4.19), (4.20), (4.21). For definiteness, we will use the last. If $\|AF\|\leqq\varepsilon$ and $\|BF\|\leqq\beta$, then

(4.40) $$\int e^{-2\tau\omega^2}|F|^2\,d\omega\leqq\varepsilon^2, \qquad \int e^{2\sigma\omega^2}|F|^2\,d\omega\leqq\beta^2=\lambda^{-2}\varepsilon^2.$$

But for all ω

(4.41) $$\min\,(e^{-\tau\omega^2},\lambda\, e^{\sigma\omega^2})\geqq e^{-\tau\omega_c^2},$$

where ω_c solves (4.39):

(4.42) $$\omega_c=[(\sigma+\tau)^{-1}\ln\lambda^{-1}]^{1/2}.$$

Therefore, by the definition of $\mathcal{M}(\varepsilon,\beta)$,

(4.43) $$\|F\|\leqq e^{\tau\omega_c^2}\varepsilon=\mathcal{M}(\varepsilon,\beta).$$

By (4.42), this says

(4.44) $$\mathcal{M}(\varepsilon,\beta)=\beta^{\tau/(\sigma+\tau)}\varepsilon^{\sigma/(\sigma+\tau)}.$$

And now (4.21) gives the error estimate

(4.45) $$\|U_c-U_0\|\leqq\sqrt{2}\beta^{\tau/(\sigma+\tau)}\varepsilon^{\sigma/(\sigma+\tau)}.$$

This upper bound pertains to the transforms, but it applies to $u_c(z)-u_0(z)$ since, by Parseval's theorem, $\|u_c-u_0\|=\|U_c-U_0\|$. From the logarithmic convexity of solutions of the heat equation, an upper bound like (4.45) is the most we could expect from any numerical method. See, for instance, [1].

For a numerical example, suppose

(4.46) $$\sigma=1,\quad \tau=1,\quad \beta=1,\quad \varepsilon=10^{-4}.$$

Then (4.45) says $\|u_c-u_0\|\leqq\sqrt{2}\ 10^{-2}$. The approximate solution is easy to compute numerically; it is the finite integral

(4.47) $$u_c(z)=\frac{1}{\sqrt{2\pi}}\int_{-\omega_c}^{\omega_c} e^{-iz\omega}\,e^{\omega^2}H(\omega)\,d\omega,$$

where $\omega_c=2.146$. Since $H(\omega)$ is the transform of the real-valued function $h(z)$, we have $H(-\omega)=\bar{H}(\omega)$, and (4.47) becomes

(4.48) $$u_c(z)=\sqrt{\frac{2}{\pi}}\int_0^{2.146} e^{\omega^2}\,\mathrm{Re}\,(e^{-iz\omega}H(\omega))\,d\omega,$$

which can easily be integrated numerically.

For deconvolution in general, the approximate solutions $u_1(z)$ and $\hat{u}(z)$ have simple analytic forms, but the cutoff solution $u_c(z)$ usually has an advantage for numerical analysis: it is an integral over a *finite* interval.

5. Ill-conditioned matrices. If A is positive definite but ill-conditioned, and if B is positive definite, and if the unknown x^0 satisfies

$$\|Ax^0-h\| \leqq \varepsilon, \qquad \|Bx^0\| \leqq \beta, \tag{5.1}$$

then the different minimum principles (1.6), (1.13) give the different approximate solutions

$$x^1=(A^2+\lambda^2B^2)^{-1}Ah, \qquad \hat{x}=(A+\lambda B)^{-1}h, \tag{5.2}$$

where $\lambda = \varepsilon/\beta$. Error estimates appear in § 2.

As $\varepsilon \to 0$, the approximation $\hat{x}$ has two numerical advantages: 1) it can be computed more quickly; 2) it can be computed more accurately, since the condition numbers of $A^2+\lambda^2B^2$ and $A+\lambda B$ must approach those of A^2 and A as $\lambda \to 0$. Thus, if $\gamma(A)$ is the condition number of A, as $\varepsilon \to 0$

$$\gamma(A^2+\lambda^2B^2) \to \gamma(A^2) = [\gamma(A)]^2 > \gamma(A), \tag{5.3}$$

while

$$\gamma(A+\lambda B) \to \gamma(A). \tag{5.4}$$

Before the limit $\lambda = 0$, the condition numbers are hard to estimate unless A and B commute. But if A and B commute, and if A and λB have the corresponding eigenvalues ρ_v and σ_v, then of course

$$\rho_v^2+\sigma_v^2 \leqq (\rho_v+\sigma_v)^2 \leqq 2(\rho_v^2+\sigma_v^2).$$

So the condition numbers of $A+\lambda B$ and $A^2+\lambda^2B^2$ satisfy

$$\frac{1}{2} \leqq \frac{[\gamma(A+\lambda B)]^2}{\gamma(A^2+\lambda^2B^2)} \leqq 2. \tag{5.5}$$

REFERENCES

[1] A. DOLD AND B. ECKMANN, EDS., *Symposium on Non-Well-Posed Problems and Logarithmic Convexity*, Springer-Verlag, New York, 1973.

[2] J. N. FRANKLIN, *Stability of bounded solutions of linear functional equations*, Math. Comput., 25 (1971), pp. 413–424.

[3] ———, *On Tikhonov's method for ill-posed problems*, Ibid., 28 (1974), pp. 889-907.

[4] A. E. HOERL, *Optimum solution of many variable equations*, Chemical Engrg. Progr., 55 (1959), pp. 69–78.

[5] ———, *Application of ridge analysis to regression problems*, Ibid., 58 (1962), pp. 54–59.

[6] ———, *Ridge analysis*, Ibid., 60 (1964), pp. 67–78.

[7] A. E. HOERL AND R. W. KENNARD, *Ridge regression: Biased estimation for nonorthogonal problems*, Technometrics, 12 (1970), pp. 55–67.

[8] ———, *Ridge regression: Applications to nonorthogonal problems*, Ibid., 12 (1970), pp. 69–82.

[9] F. JOHN, *Continuous dependence on data for solutions of partial differential equations with a prescribed bound*, Comm. Pure Appl. Math., 13 (1960), pp. 551–585.

[10] C. L. LAWSON AND R. J. HANSON, *Solving Least Squares Problems*, Prentice-Hall, Englewood Cliffs, NJ, 1974.

[11] M. M. LAVRENTIEV, *Some Improperly Posed Problems of Mathematical Physics*, Springer-Verlag, New York, 1967.

[12] D. W. MARQUARDT, *An algorithm for least-squares estimation of nonlinear parameters*, J. Soc. Indust. Appl. Math., 11 (1963), pp. 431–441.

[13] ———, *Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation*, Technometrics, 12 (1970), pp. 591–612.

[14] K. MILLER, *Least-squares methods for ill-posed problems with a prescribed bound*, this Journal, 1 (1970), pp. 52–74.

[15] D. D. MORRISON, *Methods for nonlinear least squares problems and convergence proofs*, Proc. of Seminar on Tracking Programs and Orbit Determination, Cochairmen: Jack Lorell and F. Yagi, Jet Propulsion Lab., Pasadena, CA, 1960, pp. 1–9.

[16] C. PUCCI, *Sui problemi di Cauchy non "ben posti"*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur., 18 (1955), pp. 709–722.

[17] A. N. TIKHONOV, *On the solution of incorrectly formulated problems and the regularization method*, Dokl. Akad. Nauk SSSR, 151 (1963), pp. 501–504.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0008 $01.00/0

COMPUTATIONALLY USEFUL BOUNDS FOR SINGULAR NONLINEAR SECOND ORDER BOUNDARY VALUE PROBLEMS*

LOUIS B. BUSHARD†

Abstract. Bounds are found for a solution y, to the ordinary differential equation $(r^k p(r, y)y')' + r^k q(r, y) = 0$ $(' = d/dr)$, with initial conditions $y(0) = y_0 > 0$, $y'(0) = 0$. The principal assumptions are that $k \geqq 1$, $q(r, y(r))$ is decreasing for $0 \leqq r \leqq a$ and $q(a, y(a)) \geqq 0$. Also, r-independent bounds on p and q are assumed to exist. Lower bounds on y valid for $0 \leqq r \leqq a$ are found, and, also, upper bounds valid on a possibly smaller interval. They will be applied to Bessel's equation to show how close they are. The bounds will further be applied to boundary value problems arising in gas glow discharge theory and tubular chemical reactor theory.

1. Introduction. A comparison theorem for singular nonlinear second order ordinary differential equations will be established. It will be applied to a boundary value problem arising in tubular chemical reactor theory considered by Parter and others [12]. Namely,

$$(xu')' + \beta x \exp(-1/|u|) = 0 \qquad (' = d/dx), \tag{1}$$

$$u'(0) = 0, \qquad u(1) = \tau, \tag{2}$$

where $\beta \geqq 0$ and u is positive on (0, 1). The comparison theorem will show that if u is a solution of (1) with $u(0) = u_0 > 0$, $u'(0) = 0$, then the solution $u_1(x, u_0, \beta)$ of

$$u'' + (\beta/2) \exp(-1/|u|) = 0, \tag{3}$$

$$u(0) = u_0, \qquad u'(0) = 0, \tag{4}$$

satisfies $u_1(x, u_0, \beta) < u(x)$, as long as $u(x) \geqq 0$. Further, it will be shown that (i) $u_1(x, u_0, \beta)$ is strictly monotone decreasing with respect to β, (ii) the equation $u_1(1, u_0, \beta) - \tau = 0$ has a unique C^1 solution, $\beta = \beta_*(u_0, \tau)$, $\tau \leqq u_0 < \infty$, and (iii) if u is a solution to the system (1)–(2), then either $\beta = 0$ or $\beta > 0$ and $\beta > \beta_*(u(0), \tau)$. An upper bound $\beta = \beta^*(u_0, \tau)$ is also found but it is not defined for all $u_0 \geqq \tau$. A particular value of the bounds β_* and β^* is that they significantly reduce the set of initial values that one needs to consider in a numerical shooting method of solution to the system (1)–(2). The reduction can be quite useful when it is not known how many, if any, solutions exist to systems like (1)–(2).

An example from a gas glow discharge theory will also be treated. Further, the theorem will be applied to the equation $(x^k y')' + x^k y = 0$ for $k = 1, 2$ and the results will give a measure of how close the bounds are. As pointed out by Russell and Shampine [13], differential equations of the type $(x^k u')' + x^k f(u) = 0$, $k = 0, 1, 2$, arise quite naturally and frequently in physical situations.

Existence theorems for boundary value problems based on contraction mapping, Brouwer or Schauder fixed point theorems as in [11] require estimates on solutions with respect to initial values or estimates on solution paths in appropriate topological spaces. Adequate bounds for this purpose for the gas glow discharge problem were not found in the literature. The comparison theorem arose out of an attempt to develop such bounds. The proof of the theorem relies heavily on the singularity in the differential equation. Results of a similar nature for the nonsingular problem can be found in [3] and [9]. Results for the nonsingular problem using Lyapunov methods

* Received by the editors April 2, 1976, and in revised form November 12, 1976.

† Mathematics Section, Babcock & Wilcox Research Center, Alliance, Ohio. Now at Cray Research, Inc., Chippewa Falls, Wisconsin 54729.

and Nagumo conditions can be found in [2], [7] and [8]. Many bounds are linear in nature as in [1] and [4]. Finally, [1], [10], [11] and [14] contain extensive bibliographies on two point boundary value problems as well as oscillation, comparison and disconjugacy results.

2. The comparison theorem.

THEOREM. *Let $a > 0$, $p(r, y) \in C^1$, $q(r, y) \in C^0$ where $p > 0$ and let $k \geqq 1$. Further, let $p_l(y)$, $p_u(y) \in C^1$, $q_u(y) \in C^0$ where $0 < p_l(y) \leqq p(r, y) \leqq p_u(y)$, $q(r, y) \leqq q_u(y)$ for $0 \leqq r \leqq a$. If $y(r) \in C^2$ satisfies*

$$(r^k p(r, y(r))y'(r))' + r^k q(r, y(r)) = 0 \qquad (' = d/dr) \tag{5}$$

for $0 \leqq r \leqq a$, with $y(0) = y_0 > 0$, $y'(0) = 0$, $q(a, y(a)) \geqq 0$ and $q(r, y(r))$ decreasing for $0 \leqq r \leqq a$, then $y'(r) \leqq 0$ and

$$\int_{y(r)}^{y_0} \frac{p_l(s)\, ds}{\left(\int_s^{y_0} p_u(x) q_u(x)\, dx\right)^{1/2}} \leqq r \tag{6}$$

for $0 \leqq r \leqq a$. Moreover, if $p(r, y) = p(y)$, if there is a $q_l(y) \in C^0$ such that $q_l(y) \leqq q(r, y)$, $0 \leqq r \leqq a$, if $1 \leqq k < 2$, and if $q_l(y_0) - (k/2)q_u(y_0) > 0$, then

$$\int_{y(r)}^{y_0} \frac{p(s)\, ds}{\left(2 \int_s^{y_0} p(x)(q_l(x) - Q(x))\, dx\right)^{1/2}} \geqq r \tag{7}$$

for $0 \leqq r \leqq b$, where

$$Q(y) = k \frac{\left(\int_y^{y_0} p(x) q_u(x)\, dx\right)^{1/2}}{\int_y^{y_0} \{p(s)\, ds / \left(\int_s^{y_0} p(x) q_u(x)\, dx\right)^{1/2}\}},$$

where $y(b) = z_0$, $0 < b \leqq a$, and $z = z_0$ is the largest zero smaller than y_0 of

$$\int_z^{y_0} p(x)(q_l(x) - Q(x))\, dx$$

Finally, (7) also holds for $k \geqq 1$ with $Q(y)$ replaced by $Q(y) = (k/(k+1))q(0, y_0)$ while (6) holds with $q_u(y)$ replaced by $2q_u(y)/(k+1)$.

The hypothesis that $q(r, y(r))$ is descreasing is not stringent. For example, if $q = q(y)$ with $q(0) = 0$, $y_0 > 0$, q increasing with respect to y, then $y'(r) \leqq 0$, as an integration shows, and $q(y(r))$ is decreasing for $0 \leqq r \leqq r_0$ where r_0 is the first zero of $y(r)$.

$Q(y_0)$ in the first instance of the theorem is taken to be

$$\lim_{y \to y_0} Q(y) = (k/2) q_u(y_0).$$

Equality in (6) defines a lower bound y_1 for the solution y of (5) on the interval $0 \leqq r \leqq a$, and y_1 satisfies the differential equation

$$(p_l(y_1)y_1')' + \tfrac{1}{2} p_u(y_1) q_u(y_1)/p_l(y_1) = 0 \tag{8a}$$

with the initial values of y, i.e., $y_1(0) = y_0$, $y_1'(0) = 0$. Equality in (7) defines an upper bound y_2 on $0 \leqq r \leqq b$, which satisfies

$$(p(y_2)y_2')' + q_l(y_2) - Q(y_2) = 0, \tag{8b}$$

with the initial values of y. Equality in the inequalities described at the end of the theorem define lower and upper bounds, y_3 and y_4 respectively, which satisfy

$$(p(y_3)y_3')' + (1/(k+1)) q_u(y_3) = 0 \tag{8c}$$

and

$$(p(y_4)y_4')' + q_l(y_4) - (k/(k+1))q(0, y_0) = 0, \tag{8d}$$

again, having the initial values of y. y_3 and y_4 are bounds on the intervals $0 \leqq r \leqq a$ and $0 \leqq r \leqq b$ where, in this instance, b is determined by the second choice of $Q(y) = (k/(k+1))q(0, y_0)$.

It is interesting to compare the values of $y''(0)$, $y_i''(0)$, $i = 1, 2, 3, 4$. They are

$$y''(0) = -\frac{1}{k+1}\frac{q(0, y_0)}{p(0, y_0)}, \qquad y_1''(0) = -\frac{1}{2}\frac{p_u(y_0)q_u(y_0)}{p_l^2(y_0)},$$

$$y_2''(0) = -\frac{1}{p(y_0)}\left(q_l(y_0) - \frac{k}{2}q_u(y_0)\right), \qquad y_3''(0) = -\frac{1}{k+1}\frac{q_u(y_0)}{p(y_0)},$$

$$y_4''(0) = -\frac{1}{p(y_0)}\left(q_l(y_0) - \frac{k}{k+1}q(0, y_0)\right).$$

In case $p = p_l = p_u$ and $q = q_l = q_u$, they become

$$y_1''(0) = -\frac{1}{2}\frac{q(y_0)}{p(y_0)}, \qquad y_2''(0) = -\left(1 - \frac{k}{2}\right)\frac{q(y_0)}{p(y_0)},$$

$$y''(0) = y_3''(0) = y_4''(0) = -\frac{1}{k+1}\frac{q(y_0)}{p(y_0)}.$$

When $k = 1$, all second derivatives agree at $r = 0$. When $k > 1$, $y_3''(0)$ and $y_4''(0)$ agree with $y''(0)$ and so y_3 and y_4 are closer to y for small r than are y_1 and y_2. In contrast though, y_1 is defined for a more general case of 5. When $k = 2$, y_2 is not defined. When $k = 1$, the upper bound of (8b) may be closer to y than the upper bound of (8d), as an example will show; however, the integrals in $Q(y)$ may be difficult to evaluate numerically or otherwise.

Proof. Integration of (5) gives

$$-r^k p(r, y(r))y'(r) = \int_0^r s^k q(s, y(s))\, ds, \qquad 0 \leqq r \leqq a. \tag{9}$$

Because $q(r, y(r))$ is decreasing, for $0 \leqq r \leqq a$, the inequality

$$\frac{r^{k+1}}{k+1}q(r, y(r)) \leqq -r^k p(r, y(r))y'(r) \leqq \frac{r^{k+1}}{k+1}q(0, y_0) \tag{10}$$

holds for $0 \leqq r \leqq a$. In particular, $y'(r) \leqq 0$, $0 \leqq r \leqq a$. Inequality (10) is equivalent to

$$q(r, y(r)) \leqq -(k+1)p(r, y(r))y'(r)/r \leqq q(0, y_0) \tag{11}$$

for $0 \leqq r \leqq a$. Now $-p(r, y(r))y'(r)/r$ is decreasing on $0 \leqq r \leqq a$, for

$$-(p(r, y(r))y'(r)/r)' = (q(r, y(r)) + (k+1)p(r, y(r))y'(r)/r)/r \leqq 0.$$

Returning to (9), we have

$$-r^k p(r, y(r))y'(r) = \int_0^r s^{k-1}q(s, y(s))p(s, y(s))\left(-\frac{s}{p(s, y(s))y'(s)}\right)(-y'(s))\, ds$$

$$\leqq r^{k-1}\left(-\frac{r}{p(r, y(r))y'(r)}\right)\int_0^r p_u(y(s))q_u(y(s))(-y'(s))\, ds,$$

or

$$(-p(r, y(r))y'(r))^2 \leqq \int_{y(r)}^{y_o} p_u(x)q_u(x)\,dx.$$

Finally,

$$-p_l(y(r))y'(r) \leqq \left(\int_{y(r)}^{y_0} p_u(x)q_u(x)\,ds\right)^{1/2},$$

from which (6) follows.

Next, (7) is established. The last inequality together with (6) and $p = p_u = p_l$ give

$$-k\frac{p(y(r))y'(r)}{r} \leqq k\frac{(\int_{y(r)}^{y_0} p(x)q_u(x)\,dx)^{1/2}}{\int_{y(r)}^{y_0}\{p(s)\,ds/(\int_s^{y_0} p(x)q_u(x)\,dx)^{1/2}\}}$$
$$= Q(y(r)).$$

Thus

$$\begin{aligned}&(p(y(r))y'(r))' + q_l(y(r)) - Q(y(r))\\ &\quad= (r^k p(y(r))y'(r)/r^k)' + q_l(y(r)) - Q(y(r))\\ &\quad= -q(r, y(r)) - kp(y(r))y'(r)/r + q_l(y(r)) - Q(y(r)) \leqq 0.\end{aligned}$$

Multiplying this inequality by $2p(y(r))y'(r)$ gives

$$2p(y(r))y'(r)(p(y(r))y'(r))' + 2p(y(r))(q_l(y(r)) - Q(y(r)))y'(r) \geqq 0.$$

Now $q_l(y_0) - Q(y_0) = q_l(y_0) - (k/2)q_u(y_0) > 0$ and an integration gives (7).

Next (7) is established for the case $Q(y) = (k/(k+1))q(0, y_0)$. From (11)

$$\begin{aligned}&(p(y(r))y'(r))' + q_l(y(r)) - Q(y(r))\\ &\quad= -q(r, y(r)) + q_l(y(r)) - kp(y(r))y'(r)/r - (k/(k+1))q(0, y_0) \leqq 0.\end{aligned}$$

The result follows from this inequality. The last case of (6) follows from

$$\begin{aligned}&(p(y(r))y'(r))' + (1/(k+1))q_u(y(r))\\ &\quad= -q(r, y(r)) - kp(y(r))y'(r)/r + (1/(k+1))q_u(y(r))\\ &\quad\geqq -q(r, y(r)) + (k/(k+1))q(r, y(r)) + (1/(k+1))q_u(y(r))\\ &\quad\geqq (1/(k+1))(q_u(y(r)) - q(r, y(r))) \geqq 0.\end{aligned}$$

3. Applications.

A. $(ry')' + ry = 0$, $y(0) = 1$, $y'(0) = 0$. The solution is the Bessel function of order zero, $y(r) = J_0(r)$. This application of the theorem and the next will give a measure of how close the bounds are. The bounds of (8a)–(8d) satisfy

$$y_i'' + \tfrac{1}{2}y_i = 0, \qquad i = 1, 3,$$
$$y_2'' + y_2 - \frac{1}{2}\frac{\sqrt{1-y_2^2}}{\cos^{-1} y_2} = 0,$$
$$y_4'' + y_4 - \tfrac{1}{2} = 0,$$

with $y_i(0) = 1$, $y_i'(0) = 0$, $i = 1, 2, 3, 4$. Thus $y_i(r) = \cos(r/\sqrt{2})$, $i = 1, 3$, and $y_4(r) = (1 + \cos r)/2$. y_2 is computed numerically. The results are shown in Fig. 1.

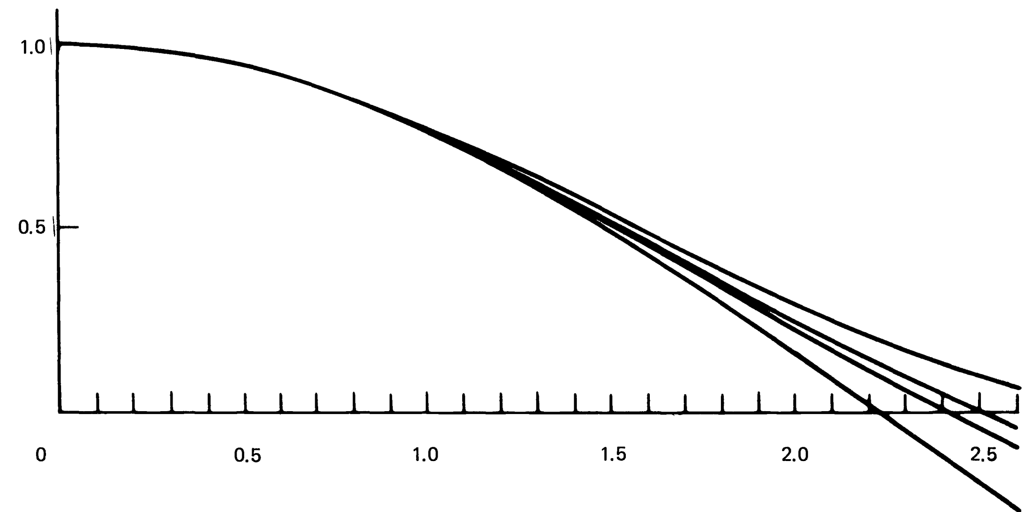

FIG. 1. y_1 ($=y_3$), J_0, y_2 *and* y_4 *respectively for application* A (*drawn to scale*).

B. $(r^2y')'+r^2y=0$, $y(0)=1$, $y'(0)=0$. The solution is $y(r)=(\sin r)/r$. $y_1=\cos(r/\sqrt{2})$ as in application A. y_2 is not defined, y_3 and y_4 satisfy

$$y_3''+\tfrac{1}{3}y_3=0, \qquad y_4''+y_4-\tfrac{2}{3}=0,$$

with $y_i(0)=1$, $y_i'(0)=0$, $i=3,4$. $y_3(r)=\cos(r/\sqrt{3})$ and $y_4(r)=(2+\cos r)/3$. For this application y_3 is better than y_1. The first zero of y is $a=\pi$ and

$$y_1(\pi)=-.606<y_3(\pi)=-.241<y(\pi)=0<y_4(\pi)=\tfrac{1}{3}.$$

Also, the first zero of y_3 is $(\sqrt{3/2})\pi=.866a$.

C. *A problem in tubular chemical reactor theory.*

$$(ry')'+\beta rf(y)=0, \tag{12}$$

$$y'(0)=0, \qquad y(1)=\tau. \tag{13}$$

In (12) and (13), f is positive for y positive and increasing, $\beta\geqq 0$, $\tau\geqq 0$ and only nonnegative solutions are of interest. The case $f=f_0$, $f_0(y)=\exp(-1/|y|)$ was briefly discussed in the Introduction. This boundary value problem was extensively analyzed mathematically by Parter [12] and it has nonunique solutions. Parter reports numerical results obtained by others for $f=f_0$ that are well supported by his mathematical analysis. Specifically, two monotone decreasing curves were computed, $\beta=\boldsymbol{\beta}(\tau)$, $\beta=\tilde{\beta}(\tau)$, $0\leqq\tau\leqq .24120$. $\boldsymbol{\beta}(\tau)<\tilde{\beta}(\tau)$ except for $\tau=.24120$ where $\boldsymbol{\beta}(\tau)=\tilde{\beta}(\tau)=10.961$. On these two curves two solutions were found, while between them three solutions were found, while in the remaining portion of the first quadrant of the τ-β plane that was sampled only one solution was found. These results were found by simply computing an approximation to $y(1,y_0,\beta)$ where $y(r,y_0,\beta)$ is the unique solution to (12) with $y(0)=y_0$, $y'(0)=0$ when $f=f_0$.

Turning to the bounds in the theorem, we note that y_1 and y_4 satisfy

$$y_1''+\tfrac{1}{2}\beta f(y_1)=0, \qquad y_4''+\beta(f(y_4)-\tfrac{1}{2}f(y_0))=0,$$

with $y_i(0)=y_0$, $y_i'(0)=0$, $i=1,4$. $y_3=y_1$ and y_2 will not be used. y_1 and y_4 are uniquely determined functions of r, y_0 and β, and they are C^1 in y_0 and β if f is C^1 in y. As long as $y(r,y_0,\beta)\geqq 0$, the inequality $y_1(r,y_0,\beta)<y(r,y_0,\beta)$ holds for $r>0$.

Also, the inequality $y(r, y_0, \beta) < y_4(r, y_0, \beta)$ holds for $r > 0$ as long as $y(r, y_0, \beta) \geqq z_0$ where z_0 is defined in the theorem.

The solution of $F_i(y_0, \beta, \tau) = y_i(1, y_0, \beta) - \tau = 0$, $i = 1, 4$, and $F(y_0, \beta, \tau) = y(1, y_0, \beta) - \tau = 0$ are needed for $\tau > 0$. In order to establish these solutions, the special solutions $y_1(r, y_0, 0) = y(r, y_0, 0) = y_4(r, y_0, 0) = y_0$ are needed. As a consequence, $y_1(1, \tau, 0) - \tau = y_4(1, \tau, 0) - \tau = y(1, \tau, 0) - \tau = 0$ and $y_{1y_0}(1, y_0, 0) = y_{4y_0}(1, y_0, 0) = y_{y_0}(1, y_0, 0) = 1$. Further, $y_{1\beta}(1, y_0, 0) = y_{4\beta}(1, y_0, 0) = y_\beta(1, y_0, 0) = -f(y_0)/4 < 0$ for $y_0 > 0$. Thus each equation has a solution. That is, there exist three C^1 functions, β_*, $\bar{\beta}$, β^*, of y_0 and τ, defined for $|y_0 - \tau|$ sufficiently small such that $y_1(1, y_0, \beta_*(y_0, \tau)) - \tau = y_4(1, y_0, \beta^*(y_0, \tau)) - \tau = y(1, y_0, \bar{\beta}(y_0, \tau)) - \tau = 0$, $\beta_*(\tau, \tau) = \beta^*(\tau, \tau) = \bar{\beta}(\tau, \tau) = 0$, and $\beta_{*\tau}(\tau, \tau) = \beta^*_\tau(\tau, \tau) = \bar{\beta}_\tau(\tau, \tau) = 4/f(y_0)$. It will be shown that β_* is defined for all $y_0 \geqq \tau$ and that $\beta = \beta_*(y_0, \tau)$ is the only solution of $y_1(1, y_0, \beta) - \tau = 0$ for $y_0 \geqq \tau$. The same statements will be found to hold for β^* on a possibly smaller y_0 set of the form $\tau \leqq y_0 \leqq y_0^*(\tau)$. Similar statements have not been found for $\bar{\beta}$, and therein lies the principal value of the application of the theorem to this example.

First, $F_1(y_0, 0, \tau) = y_0 - \tau$ shows that $F_1(y_0, 0, \tau) > 0$ for $y_0 > \tau$. The quadrature formula

$$\int_{y_1(r)}^{y_0} \frac{ds}{\left(\beta \int_s^{y_0} f(t)\, dt\right)^{1/2}} = r$$

shows that y_1 and F_1 are strictly decreasing in β for positive β and nonnegative solutions of (12) and that $F_1 < 0$ for β sufficiently large. Thus, there is exactly one $\beta = \beta_*(y_0, \tau) > 0$ such that $F_1(y_0, \beta, \tau) = 0$. The formula

$$y_{1\beta}(r) = y_1'(r) \int_0^r \frac{1}{2y_1'(t)} \int_{y_1(t)}^{y_0} f(s)\, ds\, dr$$

shows that $F_{1\beta}(y_0, \beta_*(y_0, \tau), \tau) < 0$ and, thus, β_* is C^1.

The same methods give the same results for y_4, F_4 and β^* on $\tau \leqq y_0 \leqq y_0^*(\tau)$ where $y_0^*(\tau)$ is the smallest zero $z_0 > \tau$ of

$$\int_\tau^z (f(s) - f(z)/2)\, ds.$$

Finally, if $y_0 > \tau$ and $0 < \beta \leqq \beta_*(y_0, \tau)$, then $F_{y_0}(y_0, \beta, \tau) \geqq 0$ and, consequently, $y(1, y_0, \beta) - \tau > 0$. Hence, any solution of (12)–(13) with $\tau > 0$ and $\beta > 0$ must satisfy $y_0 > \tau$ and $\beta > \beta_*(y_0, \tau)$. If $y_0 < y_0^*(\tau)$ also, then $\beta < \beta^*(y_0, \tau)$.

β_*, $\bar{\beta}$ and β^* were computed for $f = f_0$, $\tau = .05$ and $\tau \leqq y_0 \leqq 3.15$ and certain of their values are given in Table 1. The value computed for y_0^* (.05) is .0545. The qualitative nature of the computed functions β_*, $\bar{\beta}$ and β^* is depicted in Fig. 2 and the computations suggest that β_* and $\bar{\beta}$ have one local maximum and one local minimum for $y_0 > \tau$, much like a cubic polynomial. If, in fact, such behavior holds true to $\bar{\beta}$, it would provide a nice explanation of the results on the number of solutions stated by Parter which were mentioned in the introduction. The values in Table 1 show that β_* not only agrees qualitatively with $\bar{\beta}$ but also quantitatively. For example, at $y_0 = .0539$, corresponding to the local maximum of $\bar{\beta}$, $\bar{\beta}/\beta_* = 1.15$.

β_* and β^* can be found explicitly and, e.g.,

$$\bar{\beta}(y_0, \tau) > \beta_*(y_0, \tau) = \left(\int_\tau^{y_0} \frac{ds}{\left(\int_s^{y_0} f(t)\, dt\right)^{1/2}} \right)^2.$$

TABLE 1

y_0	β_*	$\bar{\beta}$	β^*
.05	0	0	0
.0501	1.88(10^5)	1.88(10^5)	1.89(10^5)
.051	1.40(10^6)	1.45(10^6)	1.51(10^6)
.052	2.03(10^6)	2.18(10^6)	2.41(10^6)
.053	2.24(10^6)	2.50(10^6)	3.03(10^6)
.0533	2.25(10^6)	2.54(10^6)	3.74(10^6)
.0539	2.23(10^6)	2.570(10^6)	
.054	2.22(10^6)	2.569(10^6)	
.055	2.08(10^6)	2.55(10^6)	
.056	1.87(10^6)	2.35(10^6)	
.15	6.65(10^2)	2.04(10^3)	
1.05	13.19	16.5	
1.15	13.14	16.1	
1.25	13.18	15.9	
1.35	13.3	15.845	
1.45	13.5	15.847	
2.25	15.6	17.4	
3.15	18.8	20.3	

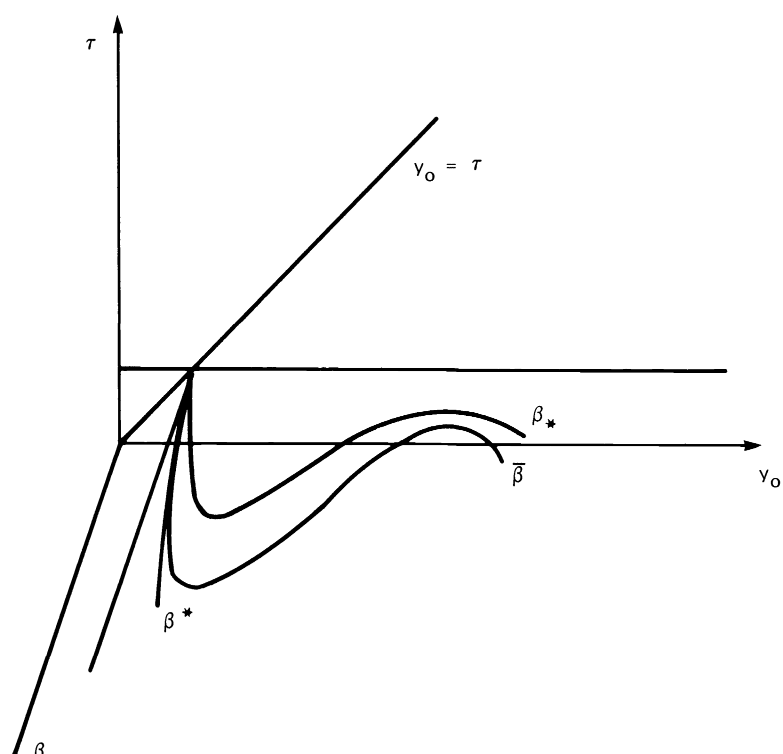

FIG. 2. *Qualitative behavior of* β_*, $\bar{\beta}$ *and* β^* *of application* C (*not drawn to scale*).

The implicit function theorem method was used above in lieu of the analytical expressions because implicit function methods extend to the case where (12) has nonlinear dependence on β.

Finally, the bound of (8b) was not used here but it could give better results than y_4 and β^*.

D. *A problem in gas glow discharge theory.*

$$(rT')' + ra(T)uT = 0, \qquad (rTu')' + r(b(T) - c(T)u)u = 0, \tag{14}$$

$$T'(0) = 0, \qquad T(1) = 1, \qquad u'(0) = 0, \qquad u(1) = 0. \tag{15}$$

In (14), $a, b, c \in C^1$ and they are positive for $T > 0$ and only solutions with $T(r)$, $u(r) > 0$, $0 \leqq r < 1$, are of physical interest. Problem (14)–(15) arises in gas glow discharge theory [5], [6]. The functions a, b and c have complicated expressions in T and will not be given—they can be found in [5] or [6]. They also depend on a parameter E_0, whose value is 400 in the present application. Several results from [6] are needed. First, the initial value problem for (14) with $T(0) = T_0$, $T'(0) = 0$, $u(0) = u_0$, $u'(0) = 0$, has unique solutions $(T(r, T_0, u), u(r, T_0, u_0))$. Secondly, $u'(r)$, $T'(r) < 0$, $0 < r \leqq 1$, for a solution to (14)–(15). Further, if $1 \leqq T_0 \leqq 2.23$, then $(a(T(r))T'(r))' \leqq 0$.

Finally, $u(r, T_0, u_0) < M(T(r, T_0, u_0), T_0, u_0)$ for $r > 0$ where $M(T, T_0, u_0)$ is defined by

$$T\frac{dM}{dT} = \max_{T \leqq s \leqq T_0} ((b(s) - c(s)M(s))/a(s)), \qquad M(T_0) = u_0.$$

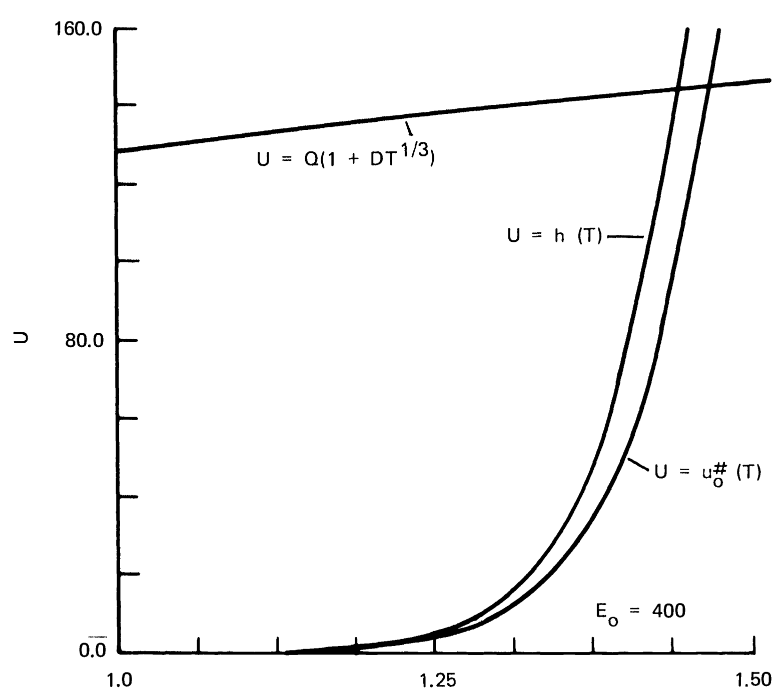

FIG. 3. *Q, h and $u_0^{\#}$ of application* D *(drawn to scale).*

M is strictly monotone increasing in u_0. The theorem applies to the first equation in (14) and it shows that the solution $T_1 = T_1(r, T_0, u_0)$ of

$$T_1'' + \tfrac{1}{2}a(T_1)T_1M(T_1, T_0, u_0) = 0$$

with $T_1(0) = T_0$, $T_1'(0) = 0$, satisfies

$$T_1(r, T_0, u_0) < T(r, T_0, u_0)$$

for $0 < r \leqq 1$. Now T_1 is strictly monotone decreasing in u_0, as can be seen from the quadrature formula for T_1. Thus, given $1 \leqq T_0 \leqq 2.23$, the equation $T_1(1, T_0, u_0) - 1 = 0$ has at most one solution $u_0 = u_0^{\#}$. Fix T_0, $1 \leqq T_0 \leqq 2.23$. If $T_1(1, T_0, \bar{u}_0) - 1 = 0$ and $T(1, T_0, u_0) - 1 = 0$, then $\bar{u}_0 < u_0$. The equation $T_1(1, T_0, u_0) - 1 = 0$ has been solved numerically for $u_0 = u_0^{\#}(T_0)$, $1 \leqq T_0 \leqq 1.48$ and the results are given in Fig. 3. Also shown in Fig. 3 are the functions $u = h(T) = b(T)/c(T)$ and $u = Q(T)$ where $u < Q(T)$ is a physical bound given in [5]. In [6], it is shown that $u_0 < h(T_0)$ for a solution to (14)–(15). Thus, a solution to (14)–(15) must have its initial values in the region enclosed by the three curves of Fig. 3. The lower bound $u_0 = u_0^{\#}(T_0)$ agrees quite closely with the lower bound found by other methods in [6]. The region of Fig. 3 was used in [5] for a thorough numerical analysis of (14)–(15).

REFERENCES

[1] P. B. Bailey, L. F. Shampine and P. E. Waltman, *Nonlinear Two-Point Boundary Value Problems*, Academic Press, New York, 1968.

[2] S. R. Bernfeld, V. Lakshmikantham and S. Leela, *Nonlinear boundary value problems and several Lyapunov functions*, J. Math. Anal. Appl., 42 (1973), pp. 545–553.

[3] L. E. Bobisud, *Comparison and oscillation theorems for nonlinear second-order differential equations and inequalities*, Ibid., 32 (1970), pp. 5–14.

[4] S. Breuer and D. Gottlieb, *Upper and lower bounds on solutions of initial value problems*, Ibid., 36 (1971), pp. 283–300.

[5] L. B. Bushard, *Mathematical and numerical analysis of a differential equation model for a gas glow discharge*, J. Computational Phys., 18 (1975), pp. 360–375.

[6] ———, *Initial value region for a boundary value problem arising in gas glow discharge theory*, SIAM J. Appl. Math., 31 (1976), pp. 547–557.

[7] R. E. Gaines, *A priori bounds for solutions to nonlinear two-point boundary value problems*, Applicable Anal., 3 (1973), pp. 157–167.

[8] J. H. George and R. J. York, *Application of Liapunov theory to boundary value problems. II*, Proc. Amer. Math. Soc., 37 (1973), pp. 207–212.

[9] R. C. Grimmer and P. Waltman, *A comparison theorem for a class of nonlinear differential inequalities*, Monatsh. Math., 72 (1968), pp. 133–136.

[10] P. Hartman, *Ordinary Differential Equations*, John Wiley, New York, 1964.

[11] L. K. Jackson, *Subfunctions and second-order ordinary differential inequalities*, Advances in Math., 2 (1968), pp. 307–363.

[12] S. V. Parter, *Solutions of a differential equation arising in chemical reactor processes*, SIAM J. Appl. Math., 26 (1974), pp. 687–716.

[13] R. D. Russell and L. F. Shampine, *Numerical methods for singular boundary value problems*, SIAM J. Numer. Anal., 12 (1975), pp. 13–36.

[14] C. A. Swanson, *Comparison and Oscillation Theory of Linear Differential Equations*, Academic Press, New York, 1968.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0009 $01.00/0

OPERATIONAL CALCULUS FOR FUNCTIONS OF TWO VARIABLES*

HARRIS S. SHULTZ†

Abstract. Let L be the space of locally integrable complex-valued functions of two variables defined on the first quadrant. We can inject L into a commutative algebra of operators on a space of testing functions. Since this injection maps convolution into multiplication it serves as a generalization (there are no growth restrictions) of the two-dimensional Laplace transform. Some useful operational formulas are developed.

Introduction. In [1] the authors develop an operational calculus for functions of two variables based on the two-dimensional Laplace transform. Many of the formulas developed therein can be obtained without using the Laplace transform and its accompanying growth restrictions. This is accomplished in [7]. In the present article we do this using an algebra of "perfect" operators on a space Q of testing functions (a two-variable analogue of [3]). This algebra contains all the functions of two variables which are locally integrable on the first quadrant as well as the operators of partial differentiation.

1. The algebra of operators.

DEFINITION 1.01. We denote by L the set of locally integrable complex-valued functions of two variables defined on the first quadrant $\{(x, y): x \geqq 0, y \geqq 0\}$ and extended to be zero elsewhere. For each f and g in L we define

$$f * g(x, y) = \int_0^y \int_0^x f(x-u, y-v)g(u, v)\,du\,dv \qquad (x \geqq 0, y \geqq 0).$$

Remarks 1.02. We may infer from (I, 2; 20), (III, 2; 2) and (III, 2; 43) of [4] that $f * g \in L$ for all f and g in L and that

$$f * g = g * f \qquad (\text{all } f, g \in L) \tag{1.03}$$

and

$$(f * g) * h = f * (g * h) \qquad (\text{all } f, g, h \in L). \tag{1.04}$$

DEFINITION 1.05. We define Q to be the subset of L consisting of those functions which are infinitely differentiable and which, along with all partial derivatives, vanish on $\{(x, y): x = 0 \text{ or } y = 0\}$.

Remark 1.06. It follows from [4, Thm. 5, p. 117] that $f * q \in Q$ for all f in L and all q in Q.

DEFINITION 1.07. A mapping A from Q into Q is called a *perfect operator* if $A(p * q) = Ap * q$ for all p and q in Q.

Remarks 1.08. If f belongs to L we denote by $\{f\}$ the mapping $q \mapsto f * q$. Thus,

$$\{f\}q = f * q \qquad (\text{all } q \in Q). \tag{1.09}$$

It follows from (1.04) that $\{f\}$ is a perfect operator. We may use [5 Thm. XIV, p. 173] and [4, Thm. 2, p. 74] to deduce that if f and g are elements of L then $\{f\} = \{g\}$ if and only if $f = g$ almost everywhere. The partial differentiation operators are denoted by D_x, D_y, D_{xx}, D_{xy}, etc. Repeated application of [2, 250] shows that each of these is a perfect operator.

Remarks 1.10. If A and B are perfect operators we denote by AB the composition of A with B; thus, $ABq = A(Bq)$ for any q in Q. Clearly, if A and B are perfect

* Received by the editors May 13, 1976, and in revised form November 2, 1976.
† Department of Mathematics, California State University, Fullerton, California 92634.

operators, then so is AB. By (1.04) and (1.09) we have

$$\{f * g\} = \{f\}\{g\} \qquad \text{(all } f, g \in L). \tag{1.11}$$

THEOREM 1.12. *The equations* $(AB)C = A(BC)$ *and* $AB = BA$ *hold for all perfect operators* A, B *and* C.

Proof. The first equation is obvious. As for the second, let $q_1, q_2, q_3, \cdots$ be a "δ-sequence" in Q (cf. [6, 1.21]). If A and B are perfect operators then, for each $q \in Q$ and for all $x, y \geqq 0$, we may use (1.03) and Definition 1.07 to deduce that

$$\begin{aligned} ABq(x, y) &= \lim_{n\to\infty} (q_n * ABq)(x, y) \\ &= \lim_{n\to\infty} AB(q_n * q)(x, y) \\ &= \lim_{n\to\infty} A(Bq_n * q)(x, y) \\ &= \lim_{n\to\infty} (Bq_n * Aq)(x, y) \\ &= \lim_{n\to\infty} (q_n * BAq)(x, y) \\ &= BAq(x, y). \end{aligned}$$

Remark 1.13. Any linear combination of perfect operators is a perfect operator.

DEFINITION 1.14. If B and X are perfect operators such that $BX = XB =$ the identity operator, then we write $X = B^{-1}$. Further, we define

$$\frac{A}{B} = AB^{-1}$$

for all perfect operators A.

DEFINITION 1.15. We denote by K the set of locally integrable complex-valued functions of a single variable which vanish on $(-\infty, 0)$.

THEOREM 1.16. *If g belongs to K and if we define*

$$\{g(x)\}q(x, y) = \int_0^x g(x-u)q(u, y)\, du \qquad (x, y \geqq 0)$$

and

$$\{g(y)\}q(x, y) = \int_0^y g(y-v)q(x, v)\, dv \qquad (x, y \geqq 0)$$

for all q in Q, then $\{g(x)\}$ and $\{g(y)\}$ are perfect operators.

Proof. Define $f(x, y) = g(x)$ for $x, y \geqq 0$ and zero elsewhere. Then,

$$\begin{aligned} \{f\}D_yq(x, y) &= \int_0^x \int_0^y f(x-u, y-v)q_y(u, v)\, dv\, du \\ &= \int_0^x \int_0^y g(x-u)q_y(u, v)\, dv\, du \\ &= \int_0^x g(x-u)\left[\int_0^y \frac{\partial}{\partial v} q(u, v)\, dv\right] du \\ &= \int_0^x g(x-u)q(u, y)\, du \end{aligned}$$

for all $x, y \geqq 0$ and all $q \in Q$; the last equality is from the fundamental theorem of calculus. Thus, $\{g(x)\} = D_y\{f\}$; it follows from Remark 1.10 that $\{g(x)\}$ is a perfect operator. The proof for $\{g(y)\}$ is similar.

2. Operational formulas.

THEOREM 2.01. *If f and f_x belong to L then*

$$\{f_x\} = D_x\{f\} - \{f(0, y)\}. \tag{2.02}$$

If f and f_y belong to L then

$$\{f_y\} = D_y\{f\} - \{f(x, 0)\}. \tag{2.03}$$

Proof. For each $q \in Q$ we may integrate by parts to obtain

$$\begin{aligned}\{f_x\}q(x, y) &= \int_0^y \int_0^x f_x(u, v)q(x-u, y-v)\,du\,dv \\ &= -\int_0^y f(0, v)q(x, y-v)\,dv + \int_0^y \int_0^x f(u, v)q_x(x-u, y-v)\,du\,dv \\ &= -\{f(0, y)\}q(x, y) + \{f\}D_x q(x, y) \\ &= [D_x\{f\} - \{f(0, y)\}]q(x, y)\end{aligned}$$

for all $x, y \geqq 0$. This proves (2.02). We may derive (2.03) in a similar manner.

THEOREM 2.04. *If g and g' belong to K then*

$$\{g'(x)\} = D_x\{g(x)\} - g(0) \tag{2.05}$$

and

$$\{g'(y)\} = D_y\{g(y)\} - g(0). \tag{2.06}$$

Proof. For each $q \in Q$ we may integrate by parts to obtain

$$\begin{aligned}\{g'(x)\}q(x, y) &= \int_0^x g'(x-u)q(u, y)\,du \\ &= -g(0)q(x, y) + \int_0^x g(x-u)q_x(u, y)\,du \\ &= [\{g(x)\}D_x - g(0)]q(x, y)\end{aligned}$$

for all $x, y \geqq 0$. This proves (2.05). We may derive (2.06) in a similar manner.

COROLLARY 2.07. *If f, f_x and f_{xy} belong to L then*

$$\{f_{xy}\} = D_y D_x\{f\} - D_y\{f(0, y)\} - D_x\{f(x, 0)\} + f(0, 0). \tag{2.08}$$

Proof. We combine (2.02), (2.03) and (2.05) to obtain

$$\begin{aligned}\{f_{xy}\} &= D_y\{f_x\} - \{f_x(x, 0)\} \\ &= D_y[D_x\{f\} - \{f(0, y)\}] - [D_x\{f(x, 0)\} - f(0, 0)] \\ &= D_y D_x\{f\} - D_y\{f(0, y)\} - D_x\{f(x, 0)\} + f(0, 0).\end{aligned}$$

LEMMA 2.09. *The perfect operator $D_x + D_y$ is invertible; its inverse is given by the equation*

$$(D_x + D_y)^{-1}q(x, y) = \int_0^\infty q(x-t, y-t)\,dt \qquad (x, y \geqq 0)$$

for all q in Q.

Proof. Define

$$Aq(x, y) = \int_0^\infty q(x-t, y-t)\, dt$$

for each $q \in Q$. By [2, 250] we have $Aq \in Q$. Further,

$$\begin{aligned} A(p * q)(x, y) &= \int_0^\infty \left[\int_0^\infty \int_0^\infty p(x-t-u, y-t-v) q(u, v)\, du\, dv\right] dt \\ &= \int_0^\infty \int_0^\infty \left[\int_0^\infty p(x-u-t, y-v-t)\, dt\right] q(u, v)\, du\, dv \\ &= Ap * q(x, y) \qquad (x, y \geqq 0) \end{aligned}$$

for all p and q in Q. Therefore, A is a perfect operator. Now,

$$\begin{aligned} A(D_x + D_y) q(x, y) &= \int_0^\infty [q_x(x-t, y-t) + q_y(x-t, y-t)]\, dt \\ &= -\int_0^\infty \frac{\partial}{\partial t}[q(x-t, y-t)]\, dt \\ &= q(x, y) \qquad (x, y \geqq 0) \end{aligned}$$

for all q in Q. Thus, $A = (D_x + D_y)^{-1}$.

THEOREM 2.10. *The equation*

$$\frac{\{f\}}{D_x + D_y} = \left\{\int_0^\infty f(x-t, y-t)\, dt\right\}$$

holds for all f in L.

Proof. For each $q \in Q$ we have

$$\begin{aligned} \left\{\int_0^\infty f(x-t, y-t)\, dt\right\} q(x, y) &= \int_0^\infty \int_0^\infty \left[\int_0^\infty f(x-u-t, y-v-t)\, dt\right] q(u, v)\, du\, dv \\ &= \int_0^\infty \left[\int_0^\infty \int_0^\infty f(x-t-u, y-t-v) q(u, v)\, du\, dv\right] dt \\ &= \int_0^\infty \{f\} q(x-t, y-t)\, dt \\ &= (D_x + D_y)^{-1}\{f\} q(x, y) \qquad (x, y \geqq 0). \end{aligned}$$

DEFINITION 2.11. If g belongs to K we define $\{g(y-x)\}$ to be the perfect operator $\{F\}$, where

$$F(x, y) = \begin{cases} g(y-x) & \text{for } x, y \geqq 0, \\ 0 & \text{otherwise,} \end{cases}$$

and we define $\{g(x-y)\}$ to be the perfect operator $\{G\}$, where

$$G(x, y) = \begin{cases} g(x-y) & \text{for } x, y \geqq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Remark 2.12. It follows from (1.09) and Definition 2.11 that

$$\{g(y-x)\}q(x,y)=\int_0^\infty\int_0^\infty g(v-u)q(x-u,y-v)\,du\,dv \qquad (x,y\geqq 0) \tag{2.13}$$

and

$$\{g(x-y)\}q(x,y)=\int_0^\infty\int_0^\infty g(u-v)q(x-u,y-v)\,du\,dv \qquad (x,y\geqq 0) \tag{2.14}$$

for all q in Q and all g in K.

THEOREM 2.15. *The equations*

$$\frac{\{g(x)\}}{D_x+D_y}=\{g(x-y)\} \tag{2.16}$$

and

$$\frac{\{g(y)\}}{D_x+D_y}=\{g(y-x\} \tag{2.17}$$

hold for all g in K.

Proof. Let $q\in Q$. From Lemma 2.09 and (2.14) we have

$$\begin{aligned}(D_x+D_y)^{-1}\{g(x)\}q(x,y)&=\int_0^\infty\int_0^\infty g(x-v-s)q(s,y-v)\,ds\,dv\\ &=\int_0^\infty\int_0^\infty g(u-v)q(x-u,y-v)\,du\,dv\\ &=\{g(x-y)\}q(x,y) \qquad\qquad (x,y\geqq 0).\end{aligned}$$

This proves (2.16). We may derive (2.17) in a similar manner.

Example 2.18. Consider the partial differential equation

$$h_x+h_y=f. \tag{2.19}$$

We use (2.02) and (2.03) to obtain

$$(D_x+D_y)\{h\}-\{h(0,y)\}-\{h(x,0)\}=\{f\}.$$

Thus,

$$\{h\}=\frac{\{f\}}{D_x+D_y}+\frac{\{h(0,y)\}}{D_x+D_y}+\frac{\{h(x,0)\}}{D_x+D_y}.$$

From Theorem 2.10, (2.16) and (2.17) it follows then that

$$\{h\}=\left\{\int_0 f(x-t,y-t)\,dt\right\}+\{h(0,y-x)\}+\{h(x-y,0)\}.$$

Combining this with Remark 1.08 and Definition 2.11 yields the solution

$$h(x,y)=\begin{cases}\displaystyle\int_0^\infty f(x-t,y-t)\,dt+h(0,y-x) & \text{for } 0\leqq x<y,\\ \displaystyle\int_0^\infty f(x-t,y-t)\,dt+h(x-y,0) & \text{for } 0\leqq y<x.\end{cases}$$

THEOREM 2.20. *If J_0 is the Bessel function of first kind of order zero then*

(2.21) $$\{J_0(2\sqrt{cxy})\} = (D_yD_x + c)^{-1} \qquad (all\ c > 0).$$

Proof. It can be verified directly that

$$\frac{\partial^2}{\partial y\ \partial x} J_0(2\sqrt{cxy}) = cJ_0''(2\sqrt{cxy}) + \frac{cJ_0'(2\sqrt{cxy})}{2\sqrt{cxy}}.$$

But $tJ_0''(t) + J_0'(t) + tJ_0(t) = 0$. Therefore,

$$\frac{\partial^2}{\partial y\ \partial x} J_0(2\sqrt{cxy}) = -cJ_0(2\sqrt{cxy}).$$

Combining this with (2.08) and the initial condition $J_0(0) = 1$, we obtain

$$-c\{J_0(2\sqrt{cxy})\} = D_yD_x\{J_0(2\sqrt{cxy})\} - D_y\{1(y)\} - D_x\{1(x)\} + 1$$

where $1(t) = 1$ for all $t \geqq 0$. Observing that $D_y\{1(y)\} = D_x\{1(x)\} = 1$ we conclude that $(D_yD_x + c)\{J_0(2\sqrt{cxy})\} = 1$.

Example 2.22. Consider the hyperbolic partial differential equation

(2.23) $$h_{xy} + ch = f \qquad (c > 0).$$

We may use (2.08) to obtain

$$(D_yD_x + c)\{h\} - D_y\{h(0, y)\} - D_x\{h(x, 0)\} + h(0, 0) = \{f\}.$$

Thus,

$$\{h\} = \frac{\{f\}}{D_yD_x + c} + \frac{D_x\{h(x, 0)\}}{D_yD_x + c} + \frac{D_y\{h(0, y)\}}{D_yD_x + c} - \frac{h(0, 0)}{D_yD_x + c}.$$

But, by (1.11) and (2.21) we have

(2.24) $$\frac{\{f\}}{D_yD_x + c} = \left\{ \int_0^y \int_0^x f(x-u, y-v) J_0(2\sqrt{cuv})\, du\, dv \right\}$$

and, by the proof of Theorem 1.16, we have

(2.25) $$\begin{aligned} \frac{D_x\{h(x, 0)\}}{D_yD_x + c} &= D_xD_y \left\{ \int_0^y \int_0^x h(x-u, 0) J_0(2\sqrt{cuv})\, du\, dv \right\} \\ &= \left\{ \frac{\partial}{\partial x} \int_0^x h(x-u, 0) J_0(2\sqrt{cuy})\, du \right\}. \end{aligned}$$

Similarly,

(2.26) $$\frac{D_y\{h(0, y)\}}{DgyD_x + c} = \left\{ \frac{\partial}{\partial y} \int_0^y h(0, y-v) J_0(2\sqrt{cxv})\, dv \right\}.$$

And, finally,

(2.27) $$-\frac{h(0, 0)}{D_yD_x + c} = \{-h(0, 0) J_0(2\sqrt{cxy})\}.$$

The solution $h(x, y)$ to (2.23) is thus the sum of the four functions which define the perfect operators in equations (2.24)–(2.27).

Remarks 2.28. This approach to an operational calculus is more general than that found in [1] but less general than that found in [7]. The space of perfect operators is isomorphic to the space of distributions on R^2 having support in the first quadrant (cf. [6, 2.18]); accordingly, questions concerning convergence of operators can be handled with little difficulty (see 2.16 and 2.19 in [6]).

REFERENCES

[1] V. A. DITKIN AND A. P. PRUDNIKOV, *Operational Calculus in Two Variables and its Applications*, Pergamon Press, Oxford, 1962.
[2] H. KESTELMAN, *Modern Theories of Integration*, 2nd ed., Dover, New York, 1960.
[3] G. KRABBE, *Operational Calculus*, Springer-Verlag, New York, 1970.
[4] L. SCHWARTZ, *Mathematics for the Physical Sciences*, Addison-Wesley, Reading, MA, 1966.
[5] L. SCHWARTZ, *Théorie des Distributions*, Hermann, Paris, 1966.
[6] H. SHULTZ, *Linear operators and operational calculus, Part II*, Studia Math., 41 (1972), pp. 1–25.
[7] S. VASILACH, *Sur un calcul opérationnel algébrique pour fonctions de deux variables*, Rev. Math. Pures et Appl., II (1957), pp. 181–238.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0010 $01.00/0

DUAL ORTHOGONAL SERIES WITH MODIFIER TENDING TO ZERO*

ROBERT P. FEINERMAN†

Abstract. In this paper we consider the basis property for $\{P\varphi_n + c_n Q\varphi_n\}$ where $\{c_n\}$ is a positive sequence converging to 0 and (as in [Feinerman and Kelman, this Journal, 1974]), $\{\varphi_n\}$ is a complete orthonormal sequence in a Hilbert space H, and P and Q are orthogonal projections on H. In [Feinerman and Kelman] it was proven that if $\{c_n\}$ converges to a positive limit then $\{P\varphi_n + c_n Q\varphi_n\}$ is an l^2 basis while in this paper we prove that if c_n converges to 0 it is not an l^2 basis. We also include an application of the result to a problem in heat transfer.

Introduction. In some recent papers we have introduced an abstract approach to dual orthogonal series. We defined a dual orthogonal series problem in an abstract Hilbert space which was a generalization into which could be fit practically all the previously individually studied cases of dual series. We have proven theorems about the completeness and basis properties (see especially [2] and [3]). In this paper we continue our study of the basis (or expansion) properties and apply the results to an example from the theory of heat transfer.

Notation. Throughout this paper we will have:

1) H is a real separable Hilbert space.

2) **P** and **Q** are subspaces of H which are orthogonal complements.

3) P and Q are the projection operators from H onto **P** and **Q** respectively (so that $P+Q$ is the identity operator).

4) $\{\varphi_n\}_{n=1}^{\infty}$ is a complete orthonormal sequence in H.

5) For $n = 1, 2, \cdots$ we consider $a_n P\varphi_n + b_n Q\varphi_n$ where $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ are sequences of nonnegative real numbers, one of which (say $\{a_n\}$) is a positive sequence. Since a_n is never zero, we will divide by a_n and let ψ_n be defined by $\psi_n = P\varphi_n + c_n Q\varphi_n$ where $c_n = b_n/a_n$. $\{\varphi_n\}$ is called the kernel of the dual orthogonal series and $\{c_n\}$ is called the modifier.

The dual orthogonal series problem (as defined in [3]) is: Given $f \in H$, find $\{k_n\}_{n=1}^{\infty} \in l^2$ such that $\sum_{n=1}^{\infty} k_n \psi_n = f$ (where convergence is in the norm of H). The main result of [3] can be stated as:

THEOREM 1. *If $\{c_n\}_{n=1}^{\infty}$ is a positive sequence which converges to a positive limit, then $\{\psi_n\}_{n=1}^{\infty}$ is an l^2 basis in H; i.e., for each $f \in H$ there is a unique sequence $\{k_n\}_{n=1}^{\infty}$ in l^2 such that $\sum_{n=1}^{\infty} k_n\psi_n = f$.*

In this paper we concern ourselves with the case when $\{c_n\}_{n=1}^{\infty}$ is a sequence which converges to zero. We prove that in that case $\{\psi_n\}_{n=1}^{\infty}$ is not an l^2 basis; i.e., it is not true that for each $f \in H$ there is a unique sequence $\{k_n\}_{n=1}^{\infty}$ in l^2 such that $\sum_{n=1}^{\infty} k_n\psi_n = f$. (Actually in our theorem when we prove that $\{\psi_n\}$ is not an l^2 basis if $\{c_n\}$ converges to zero, we have one other minor hypothesis. However, that hypothesis is satisfied by every dual orthogonal series problem we have encountered in the literature.)

Note. If, instead of $\{P\varphi_n + c_n Q\varphi_n\}$, we were considering $\{a_n P\varphi_n + b_n Q\varphi_n\}$ [as described in notation 5)], $\{k_n\}_{n=1}^{\infty}$ in l^2 would obviously be replaced by $\{k_n a_n\}$ in l^2.

The type of boundary problem in which one has $\{c_n\}$ converging to zero is usually one where the dual orthogonal series is associated with a mixed boundary value problem in which one of the conditions is a Dirichlet condition and the other condition is a Neumann condition. In the latter part of this paper we consider in detail a well-known problem from heat transfer which exhibits this behavior.

* Received by the editors November 18, 1975, and in revised form January 4, 1977.

† Department of Mathematics, Lehman College, City University of New York, Bronx, New York 10468.

Before we can prove our main theorem, we will need a preliminary lemma.

LEMMA. *If* $\{c_n\}_{n=1}^{\infty}$ *is a bounded sequence and* $\{\psi_n\}$ *an* l^2 *basis, then the mapping* $T\colon \sum_{n=1}^{\infty} a_n\varphi_n \to \sum_{n=1}^{\infty} a_n\psi_n$ *is an isomorphism from* H *to* H.

Proof. For any $h \in H$, we have $h = \sum_{n=1}^{\infty} a_n\varphi_n$. Then

$$\begin{aligned} \|Th\|^2 = \left\|\sum_{n=1}^{\infty} a_n\psi_n\right\|^2 &= \left\|P\sum_{n=1}^{\infty} a_n\varphi_n\right\|^2 + \left\|Q\sum_{n=1}^{\infty} a_nc_n\varphi_n\right\|^2 \\ &\leqq \left\|\sum_{n=1}^{\infty} a_n\varphi_n\right\|^2 + \left\|\sum_{n=1}^{\infty} a_nc_n\varphi_n\right\|^2 = \|h\|^2 + \sum_{n=1}^{\infty} (a_nc_n)^2 \\ &\leqq \|h\|^2 + M^2\sum_{n=1}^{\infty} a_n^2 \quad (\text{where } M = \sup_n |c_n|) \\ &= (1+M^2)\|h\|^2 \end{aligned}$$

and thus T is bounded. On the other hand, since $\{\psi_n\}$ is an l^2 basis, T is obviously one to one and onto. Hence T^{-1} exists and moreover (see [1]) T^{-1} is bounded.

THEOREM 2. *If* $\dim QH = \dim \mathbf{Q}$ *is infinite and* $\{c_n\}$ *converges to* 0, *then* $\{\psi_n\}$ *is not an* l^2 *basis.*

Proof. Set $M_n = \sup\{|c_j|; j > n\}$. Then M_n converges to zero. We set $H_n = \operatorname{sp}\{\varphi_1, \cdots, \varphi_n\}$ and let K_N be the direct sum of QH_N and PH; i.e., $K_N = QH_N \oplus PH$. Elements of K_N have the form

$$\sum_{n=1}^{N} a_nQ\varphi_n + \sum_{n=1}^{\infty} b_nP\varphi_n = \sum_{n=1}^{N} (b_nP\varphi_n + a_nQ\varphi_n) + \sum_{n=N+1}^{\infty} b_nP\varphi_n.$$

Since

$$\begin{aligned} K_N^{\perp} &= (QH_N)^{\perp} \cap (PH)^{\perp} \\ &= (QH_N)^{\perp} \cap (QH) \end{aligned}$$

we need only choose $N+1$ linearly independent elements of QH to establish that K_N is not dense in H (and since we are given that $\dim QH$ is infinite, that obviously can be done).

Assume $\{\psi_n\}_{n=1}^{\infty}$ were an l^2 basis in H. Then, for $k \in H$,

$$k = \sum_{n=1}^{\infty} a_n\psi_n = T\left(\sum_{n=1}^{\infty} a_n\varphi_n\right).$$

We define the sequence of operators $S_N\colon H \to H$ by $S_Nk = \sum_{n=1}^{N} a_n\psi_n + \sum_{n=N+1}^{\infty} a_nP\varphi_n$ and note that $S_NH \subseteq K_N$. Then

$$\begin{aligned} \|(1-S_N)k\|^2 &= \left\|\sum_{n=N+1}^{\infty} a_n\psi_n - \sum_{n=N+1}^{\infty} a_nP\varphi_n\right\|^2 \\ &= \left\|\sum_{n=N+1}^{\infty} a_nc_nQ\varphi_n\right\|^2 \leqq \left\|\sum_{n=N+1}^{\infty} a_nc_n\varphi_n\right\|^2 \\ &\leqq M_N^2\sum_{n=N+1}^{\infty} a_n^2 \leqq M_N^2\sum_{n=1}^{\infty} a_n^2 = M_N^2\|T^{-1}k\|^2 \\ &\leqq M_N^2\|T^{-1}\|^2\|k\|^2. \end{aligned}$$

Thus $\|1-S_N\| \leqq M_N\|T^{-1}\|$ which, for large enough N, is <1. Thus, for large enough N, S_N would be invertible (and hence onto) which is impossible inasmuch as $S_NH \subseteq K_N$ which is never dense.

Application. As an illustration of an application of Theorem 2, we use the following well-known problem in heat transfer theory (and obtain a result which is part of the folklore of the field).

We seek the steady temperature $U(x, y)$ in a semi-infinite rectangular strip $R=\{(x, y): 0 \leqq x \leqq \pi, y \geqq 0\}$ with zero temperature on the side walls and mixed (Dirichlet and Neumann) boundary conditions along the base. In addition, we assume $U(x, y)$ approaches zero as y approaches infinity.

The mixed boundary conditions along the base can be expressed as:

$$U_y(x, 0)=f_1(x) \quad \text{for } 0<x<c,$$
$$U(x, 0)=f_2(x) \quad \text{for } c<x<\pi$$

where $0<c<\pi$, $f_1(x) \in L^2[0, c]$ and $f_2(x) \in L^2[c, \pi]$.

By the standard method of separation of variables we get that $U(x, y)=\sum_{n=1}^{\infty} a_n e^{-ny} \sin(nx)$ where $\{a_n\}$ is given by

$$\sum_{n=1}^{\infty} a_n n \sin(nx)=-f_1(x) \quad \text{for } 0<x<c,$$

$$\sum_{n=1}^{\infty} a_n \sin(nx)=f_2(x) \quad \text{for } c<x<\pi.$$

In [4, p. 152] is established the existence of a unique solution $\{a_n\}$ to these dual series.

To fit this dual series into our abstract form, we let $H=L^2[0, \pi]$, $\mathbf{P}=L^2[0, c]$, $\mathbf{Q}=L^2[c, \pi]$ and $\varphi_n=\sqrt{(2/\pi)} \sin(nx.)$ If we then set $k_n=na_n$, $c_n=1/n$ and

$$g(x)=\begin{cases} -\sqrt{\dfrac{2}{\pi}} f_1(x), & 0<x<c, \\ \sqrt{\dfrac{2}{\pi}} f_2(x), & c<x<\pi \end{cases}$$

our dual series problem becomes: given $g \in H$, find $\{k_n\}_{n=1}^{\infty}$ such that

$$\sum_{n=1}^{\infty} k_n(P\varphi_n+c_n Q\varphi_n)=g.$$

We note that $\{c_n\}$ converges to zero and $QH=\mathbf{Q}$ has infinite dimension. Therefore, by our theorem, $\{P\varphi_n+c_n Q\varphi_n\}_{n=1}^{\infty}$ is not an l^2 basis; i.e., for some $g \in H$ (and therefore for some f_1 and f_2) the solution $\{k_n\}$ (to the abstract dual series problem) is not in l^2.

The physical significance of the solution not being in l^2 is the following. We consider the heat flux across the line at height $y=h$ and see what happens as h approaches 0. We measure the flux by considering the $L^2[0, \pi]$ integral of U_y along $y=h$. We get:

$$\int_0^{\pi} |U_y(t, h)|^2 \, dt=\int_0^{\pi}\left(\sum_{n=1}^{\infty} a_n n e^{-nh} \sin(nt)\right)^2 dt$$
$$=\sum_{n=1}^{\infty} a_n^2 n^2 e^{-2nh}=\sum_{n=1}^{\infty} k_n^2 e^{-2nh}.$$

As h approaches 0 this approaches $\sum_{n=1}^{\infty} k_n^2$. Thus, the implication of Theorem 2 is that solutions exist (with f_1, f_2 in L^2) for which the heat flux across $y=h$ gets arbitrarily large as h approaches 0.

Acknowledgment. The author would like to express his appreciation to the referee for his clever and useful suggestions.

REFERENCES

[1] G. BACHMAN AND L. NARICI, *Functional Analysis*, Academic Press, New York, 1966.

[2] R. FEINERMAN AND R. KELMAN, *The convergence of least squares approximations for dual orthogonal series*, Glasgow Math. J., 15 (March 1974), Part I, pp. 82–84.

[3] R. KELMAN AND R. FEINERMAN, *Dual orthogonal series*, this Journal, (1974), pp. 489–502.

[4] I. SNEDDON, *Mixed Boundary Value Problems in Potential Theory*, North-Holland, Amsterdam, 1966.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0011 $01.00/0

ISOPERIMETRIC INEQUALITIES IN A CLASS OF NONLINEAR EIGENVALUE PROBLEMS*

P. S. CROOKE† AND R. P. SPERB‡

Abstract. In this work we prove an isoperimetric inequality for the eigenvalue λ and other quantities in the problem $\Delta u+\lambda u^{2p+1}=0$ in D, $u=0$ on ∂D where D is a plane, bounded domain.

Introduction. The main purpose of this paper is to develop some isoperimetric and nonisoperimetric equalities connected with the nonlinear eigenvalue problem: $\Delta u+\lambda f(u)=0$ in D and $u=0$ on ∂D where D is a bounded, two dimensional domain with sufficiently smooth boundary ∂D. Using the level lines of u, a fundamental differential inequality is developed from which several of the classical isoperimetric inequalities (e.g. Faber–Krahn inequality, St. Venant principle) can be derived, along with some new results. Although the fundamental inequality is derived for general $f(u)$, the primary interest of the paper is for the special case: $f(u)=u^{2p+1}$, $p=0, 1, 2, \cdots$.

In the first section a survey of existence results for the general eigenvalue problem is presented. It is shown in the special case of $f(u)=u^{2p+1}$ that a positive eigenfunction, normalized so that its Dirichlet integral is one, exists. Also in the case that D is a disk, we prove that a radial symmetric eigenfunction exists and compute its first eigenvalues for $p=0, 1, 2, \cdots$. In the second section the fundamental inequality is derived. In § 3 the fundamental inequality is used to derive some isoperimetric inequalities. In the fourth section, other inequalities are developed using different techniques i.e., conformal mapping, a Rellich-type identity. The remarks in § 5 finally point out some directions in which the results can be extended.

1. Existence of a positive eigenfunction. Let D denote a bounded, two-dimensional region with piecewise smooth boundary ∂D. In this section we will be interested in the existence of positive solutions for the boundary-value problem:

$$(1.1)\qquad \begin{aligned} \Delta u+\lambda f(u)&=0 \quad \text{in } D,\\ u&=0 \quad \text{on } \partial D, \end{aligned}$$

where $f(u)\geqq 0$ for $u\geqq 0$ and λ is a positive real constant. Although our primary interest is in the special case when $f(u)=u^{2p+1}$, $p=0, 1, 2, \cdots$, many of the results derived in the following sections will be applicable for the general problem and hence, a summary of some of the important existence theorems for (1.1) seems to be appropriate. We will start by reviewing some results for the general problem and conclude with showing that the problem for $f(u)=u^{2p+1}$ has a positive eigenfunction under the normalization

$$\mathscr{D}(u)\equiv\int_D |\nabla u|^2\,dx=1.$$

We will also demonstrate the existence of a positive, radial-symmetric eigenfunction in the case when D is a disk. The boundary-value problem (1.1) arises in several physical situations. The reader is referred to the work of Gel'fand [9].

* Received by the editors August 19, 1976, and in revised form December 10, 1976.
†Department of Mathematics, Vanderbilt University, Nashville, Tennessee 37325.
‡Abteilung Pharmacologie, Biozentrum der Universität Basel, Basel, Switzerland.

In the existence theorems for solutions of (1.1), there are two important cases, namely, $f(0)>0$ and $f(0)=0$. In the latter case we will also assume that $f'(0)=0$. For the case $f(0)>0$ Keller and Cohen [14] have conducted an extensive study of existence results on a class of boundary value problems of which (1.1) is a special case. They have shown that provided $f(u)$ is continuous for $u\geqq 0$ and satisfies certain monotonicity requirements, then there exists a real, positive number λ^* such that if $\lambda>\lambda^*$, then there does not exist a positive solution of (1.1). The *spectrum* of (1.1) is the set of $\lambda\in\mathbb{R}^+$ such that a positive solution of (1.1) exists. They showed that the spectrum of (1.1) is either $(0, \lambda^*)$ or $(0, \lambda^*]$. Also, with f suitably restricted, for each λ in the spectrum there is a minimal (smallest) positive solution, $u(x;\lambda)$, of (1.1). The problem of computing λ^* for a given $f(u)$ and region D is, in general, difficult. It follows from the results of Keller and Cohen [14] and Laetsch [15] that for each λ in the spectrum

$$\lambda \leqq \mu_1\left\{\frac{\partial f}{\partial u}[u(x;\lambda)]\right\}, \tag{1.2}$$

where $\mu_1\{\rho(x)\}$ is the smallest eigenvalue for the inhomogeneous, fixed membrane problem:

$$\begin{aligned} &\Delta v+\mu\rho(x)v=0 \quad &&\text{in } D,\\ &v=0 &&\text{on } \partial D, \end{aligned} \tag{1.3}$$

and $u(x;\lambda)$ is the minimal positive solution for λ. Laetsch required that f be Hölder continuous on $u\geqq 0$. Bandle [4] has given several results which characterize the spectrum of (1.1) under the restrictions that f is Hölder continuous, $f'(u)>0$ and $f''(u)\geqq 0$ for $u>0$. In particular, she showed that if there exists an arbitrary, nondecreasing positive function $f_0(u)$ such that $f(u)\leqq f_0(u)$ and there exists a positive number m_0 such that $m/f_0(m)$ takes its maximum at m_0, then (1.1) will have a positive solution for each $\lambda\in\mathbb{R}^+$ which satisfies the inequality

$$\lambda \leqq \frac{4\pi m_0}{\bar{A} f_0(m_0)}, \tag{1.4}$$

where $\bar{A}$ denotes the area of D. In conjunction with this inequality and (1.2), she also showed that

$$\frac{4\pi m_0}{\bar{A} f(m_0)} < \lambda^* < \mu_1\{f'(0)\}. \tag{1.5}$$

In the case that $f(u)$ is an increasing function, Bandle and Hersch [5] have announced the following results: For a domain of given area, λ^* is a minimum for the disk. Sperb [27] has recently given a similar result which is optimal in a different sense.

Hudjaev [12] showed a necessary condition that (1.1) is solvable for any $\lambda\in\mathbb{R}^+$ is $\lim_{u\to+\infty}\inf f(u)/u=0$ and a sufficient condition is $\lim_{u\to+\infty} f(u)/u=0$. A necessary and sufficient condition that (1.1) is not solvable for every $\lambda>0$ is $\min_{u>0} f(u)/u>0$. He also showed that

$$\lambda^* \leqq \frac{\lambda_0}{\min\limits_{u>0}(f(u)/u)} \tag{1.6}$$

where λ_0 is the smallest eigenvalue of the homogeneous, fixed membrane problem for D.

In contrast to the case $f(0)>0$ where the spectrum of (1.1) might be a bounded interval of $\mathbb{R}^+$, the spectrum of (1.1) with $f(0)=0$ is $\mathbb{R}^+$, provided $f(u)$ is suitably restricted. Levinson [16] has shown that if: (1) $f(0)=0$ and $f(u)>0$ for $u>0$; (2) f satisfies a local Lipschitz condition on $\mathbb{R}^+$; (3) for $K\geqq 1$, we have

$$f(u)\log[f(u)]\leqq KF(u), \quad u>0,$$

where

$$F(u)=\int_0^u f(v)\,dv;$$

(4) for some $K_1\in(0,\frac{1}{2})$, $f(u)$ satisfies $K_1F(u)\leqq uf(u)$; then for at least one $\lambda>0$, (1.1) has a positive solution in $C^2(D)\cap C^0(\bar{D})$. Furthermore, if $\lim_{u\to 0} f(u)/u=0$ and/or $\lim_{u\to+\infty} f(u)/u=0$, then the spectrum of (1.1) has an accumulation point at $+\infty$; if $\lim_{u\to+\infty} F(u)/u^2=+\infty$, then the spectrum of (1.1) has an accumulation point at 0. For $f(u)=u^q$, $q>1$, the spectrum of (1.1) is simply $\mathbb{R}^+$. Pohozaev [23] strengthened Levinson's results and proved that if f is sufficiently smooth and satisfies the mild growth condition

$$|f(u)|\leqq A+B|u|^b e^{c|u|^\alpha}, \quad u>0,$$

where $\alpha<2$, A, B, b, c are arbitrary positive constants, and there exists a sufficiently smooth function v such that

$$\int_D F(v)\,dx\equiv\gamma\neq 0,$$

then there exists an eigenfunction $\phi\in C^2(D)\cap C^0(\bar{D})$ of (1.1) such that

$$\int_D F(\phi)\,dx=\gamma.$$

To conclude our survey we mention that Amann [1] has given a set of five theorems which contain the essence of many of the results summarized above for $f(u)>0$, $u\geqq 0$.

We now turn our attention to the problem of showing that there exists a positive eigenfunction ϕ of (1.1) with $f(u)=u^q, q>1$, such that $\mathscr{D}(\phi)=1$. Joseph and Lundgren [13] have proven the existence of radial-symmetric, positive solutions for $f(u)>0$, $u\geqq 0$, and $f'(0)>0$. For our case, the approach will be quite different. Levinson proved that the spectrum of (1.1) in this case is $\mathbb{R}^+$. The proof of this is elementary and will not be repeated here. Levinson also showed that $m(\lambda)\equiv\mathscr{D}(\phi(x;\lambda))$ is $C^0([0,+\infty))$, $m(0)=0$ and m is monotonically increasing on $[0,+\infty)$. Let $\lambda_0\in(0,+\infty)$ be fixed. We then know that there exists a positive eigenfunction ϕ_0. Set $m(\lambda_0)=m_0$. Let c be an arbitrary positive constant and define $\phi=c\phi_0$. Then ϕ satisfies $\Delta\phi+\lambda\phi^q=0$ in D, and $\phi=0$ on ∂D, with $\lambda=\lambda_0c^{1-q}$. Furthermore,

$$m(\lambda)\equiv\mathscr{D}(\phi(x;\lambda))=c^2\mathscr{D}(\phi_0(x;\lambda_0))=c^2m_0.$$

Hence, $\phi=(1/\sqrt{m_0})\phi_0$ is a solution and $\mathscr{D}(\phi)=1$ and in view of the monotonicity of $m(\lambda)$, λ is unique.

To conclude this section we prove the existence of a radial-symmetric, positive eigenfunction of (1.1) with $f(u)=u^{2p+1}$, $p=0,1,\cdots$, and D a disk. We will need such a result when we develop some of our inequalities in the next section. If D is a disk of radius R and $u=u(r)$, then (1.1) reduces to

$$ru''+u'+\lambda ru^{2p+1}=0, \quad r\in(0,R), \qquad u'(0)=0, \quad u(R)=0. \tag{1.7}$$

One can convert (1.7) into an initial-value problem (see e.g. [6], [7]) by introducing new dependent and independent variables: $z=r\sqrt{\lambda}\,[u(0)]^p$ and $u(r)=u(0)y(z)$. Then (1.7) becomes

$$zy''+y'+zy^m=0, \qquad y(0)=1, \quad y'(0)=0, \tag{1.8}$$

where we have set $m=2p+1$. It can be shown (see Bellman [29]) that $y(z)$ must oscillate. We will show that there exists a solution of (1.8) on $\mathbb{R}^+$.[1] The idea of the proof is to show first that a solution exists in a neighborhood of $z=0$ and then show that this solution can be continued to $+\infty$.

Formally, we expand $y(z)$ in a power series

$$y(z)=\sum_{n=0}^{\infty} c_n z^{2n}.$$

The coefficients are uniquely determined by the recurrence relation

$$4n^2c_n=-\left[\left(1+\sum_{r=1}^{n-1} c_r z^{2r}\right)^m\right]\Big|_{z^{2n-2}}. \tag{1.9}$$

Here we are using the notation: $[P(z)]|_{z^m}=$ the coefficient of z^m in $P(z)$. The right side of (1.9) is some polynomial, $-P(c_1, c_2, \cdots, c_{n-1})$. A straightforward calculation shows that

$$|4n^2c_n|\leqq P(|c_1|,\cdots,|c_{n-1}|)=\left[\left(1+\sum_{r=1}^{n-1}|c_r|z^{2r}\right)^m\right]\Big|_{z^{2n-2}}. \tag{1.10}$$

For $Y(z)\equiv\sum_{n=0}^{\infty}|c_n|z^{2n}$, a simple computation shows that $Y''(z)\ll Y^m(z)$ where the symbol "$\ll$" means that each coefficient in the power series for $Y''(z)$ is not greater than the corresponding coefficient in $Y^m(z)$. We now consider the following auxiliary initial-value problem:

$$W''=W^m(z), \qquad W(0)=1, \quad W'(0)=0. \tag{1.11}$$

Since this differential equation has no finite singularities, $W(z)$ is given by the power series expansion $\sum_{n=0}^{\infty} C_n z^{2n}$ where $C_0=1$ and $(2n)(2n-1)C_n=P(C_1,\cdots,C_{n-1})$ and this expansion will converge on some interval $0\leqq z<\varepsilon$. We now note that $c_n\leqq|c_n|\leqq|C_n|$ and hence, $y(z)\ll Y(z)\ll W(z)$ which implies that the power series for $y(z)$ converges at least on $0\leqq z<\varepsilon$.

Having the local existence, we now show that $y(z)$ can be extended i.e., we show that $y(z)$ and $y'(z)$ stay bounded on $0\leqq z<+\infty$. To do this we introduce new variables:

$$\xi(z)=y(z), \quad \eta(z)=zy'(z),$$

[1] The proof of this result has been most generously suggested to us by Professor Richard Arenstorf of Vanderbilt University. It is both elementary and elegant and is a model technique which can be used to prove existence for many problems of the form $zy''+y'+zf(y)=0$ such that $y(0)=1$, $y'(0)=0$.

and define a Hamiltonian $H(\xi, \eta; z) \equiv (z/(m+1))\xi^{m+1} + (1/(2z))\eta^2$ so that $\xi' = \partial H/\partial\eta$ and $\eta' = -\partial H/\partial\xi$. For $h(z) = H(\xi(z), \eta(z); z)$, we have: (i) $h(0) = 0$, (ii) $h(z) > 0$ for $0 < z < \varepsilon$; and (iii) $dh/dz \leqq (1/z)h(z)$. If z_1 is any point in $(0, \varepsilon)$, then (iii) implies that $h(z) \leqq h(z_1)z/z_1$. Since $h(0) = 0$, we have that $h(z_1)/z_1 \to 1/(m+1)$ as $z_1 \to 0$ and hence,

$$h(z) \leqq \frac{z}{m+1}. \tag{1.12}$$

Returning to the definition of $h(z)$, we note that (1.12) is equivalent to

$$\frac{1}{n+1}[y(z)]^{m+1} + \frac{1}{2}[y'(z)]^2 \leqq \frac{1}{m+1}, \quad \forall z \in (0, \varepsilon).$$

This establishes the extendibility of $y(z)$.

In the following sections we will need expressions for the first eigenvalues for (1.7). One can use the techniques of [7] to show that

$$\lambda = \frac{[\pi(p+1)]^p z_0^{2p+2}[y'(z_0)]^{2p}}{R^2}$$

where z_0 is the first positive zero of the solution of (1.8). To get an idea of the relative magnitude of λ for different values of p, (1.8) has been solved numerically to find z_0 and $y'(z_0)$. The following table summarizes these calculations.

TABLE 1

p	$R^2\lambda \doteq$
$p=0$	5.78π
$p=1$	14π
$p=2$	68π
$p=3$	235π

2. A differential inequality associated with the solution of (1.1). Let D be a simply connected plane domain with a piecewise analytic boundary ∂D. If $f(u)$ is analytic for $u \geqq 0$, then the solution of (1.1) is also analytic in D. This follows from the results of [6], [17]. Denote by $D(t)$ the subdomain of D where $u > t$ whose boundary is the level line $\Gamma(t)$ (i.e. where $u = t$). $\Gamma(t)$ does not have to be a simple curve. We mention in passing that if e.g. D is convex all level lines are simple closed curves, as was shown in [26]. Denote by $A(t)$ the area bounded by $\Gamma(t)$ and by $\bar{A}$ the total area of D. Since u is analytic we have for almost all $t \in (0, u_{\max})$

$$\frac{dA}{dt} = -\oint_{\Gamma(t)} \frac{ds}{|\nabla u|}, \quad (\nabla\text{: gradient}). \tag{2.1}$$

We introduce now the function

$$E(t) \equiv \int_t^{u_{\max}} f(v)\, dv \oint_{\Gamma(v)} \frac{ds}{|\nabla u|} = \int_{D(t)} f(u)\, dx \quad (dx = \text{area element of } D). \tag{2.2}$$

By Green's identity we have

$$\lambda E(t) = \oint_{\Gamma(t)} |\nabla u|\, ds. \tag{2.3}$$

From (2.1), (2.2) and Schwarz's inequality we get

$$-\lambda \frac{dA}{dt} E(t) \geqq \left(\oint_{\Gamma(t)} ds\right)^2, \tag{2.4}$$

and from the classical isoperimetric inequality

$$-\lambda \frac{dA}{dt} E(t) \geqq 4\pi A(t). \tag{2.5}$$

Introducing the area A as our independent variable we find from (2.2) and (2.5) finally

$$\frac{du}{dA} + \frac{\lambda E}{4\pi A} \geqq 0, \quad \text{a.e. in } (0, \bar{A}). \tag{2.6}$$

Note that (2.6) can be written in terms of $E(A)$ only. Since $dE/dA = f(u(A))$, we have

$$\frac{d^2E}{dA^2} + \frac{\lambda}{4\pi A} \frac{df}{du} E \geqq 0 \quad \text{a.e. in } (0, \bar{A}), \tag{2.7a}$$

$$E(0) = 0, \quad \frac{dE}{dA}(\bar{A}) = 0. \tag{2.7b}$$

Note also that $u(0) = u_{\max}$, $u(\bar{A}) = 0$. In the cases $f(u) = u$ and $\lambda f(u) \equiv 2$ the inequality (2.7) is linear, and in the latter, one has to replace λE by $2A$. For $f(u) = u$ (1.1) is the eigenvalue problem of the vibrating membrane spanned over D. In this case u will be taken as the first (positive) eigenfunction. For $\lambda f(u) \equiv 2$ (1.1) becomes the classical torsion problem for an elastic beam of cross-section D.

We remark that it would be possible to consider functions other than $E(t)$. In particular, for the function

$$H(t) \equiv \int_t^{u_{\max}} dv \oint_{\Gamma(v)} |\nabla u| \, ds, \tag{2.8}$$

with $H(0) = \int_D |\nabla u|^2 \, dx \equiv \mathscr{D}(u)$ one finds for the Dirichlet integral $\mathscr{D}(u)$ employing similar arguments as before

$$\mathscr{D}(u) \geqq \int_0^{\bar{A}} 4\pi A \left(\frac{du}{dA}\right)^2 dA. \tag{2.9}$$

The equality sign holds here if D is a disk and u is a function of the radius only. A number of interesting inequalities follow from (2.6). This is the subject of our next section.

3. Isoperimetric inequalities following from (2.6). Using (2.6), we give a simple proof of an extension of the well-known Faber–Krahn inequality (see e.g. [18]) to the problem which is our primary interest in this paper, namely,

$$\begin{aligned} &\Delta u + \lambda u^{2p+1} = 0 \quad \text{in } D, \qquad p = 0, 1, 2, 3, \cdots, \\ &u = 0 \quad \text{on } \partial D, \quad \mathscr{D}(u) = 1. \end{aligned} \tag{3.1}$$

As was shown in § 1 there is exactly one value of $\lambda > 0$ where a positive solution exists. Furthermore, if D is a disk, u is radially symmetric.

We multiply (2.6) by $4\pi A(du/dA)$, use the fact that $du/dA \leqq 0$ a.e. and integrate over A from 0 to $\bar{A}$ to obtain

$$\int_0^{\bar{A}} 4\pi A\left(\frac{du}{dA}\right)^2 dA \leqq \lambda \int_0^{\bar{A}} u^{2p+2}\, dA. \tag{3.2}$$

Because of the normalization of u and (2.9) we have

$$\lambda = \frac{1}{\int_D u^{2p+2}\, dx} \geqq \frac{|\int_0^{\bar{A}} 4\pi A(du/dA)^2\, dA|^{p+1}}{\int_0^{\bar{A}} u^{2p+2}\, dA} \geqq \min_{\substack{v(\bar{A})=0\\ v(0)<\infty}} \frac{|\int_0^{\bar{A}} 4\pi A(dv/dA)^2\, dA|^{p+1}}{\int_0^{\bar{A}} v^{2p+2}\, dA}. \tag{3.3}$$

Since the variational characterization of λ is

$$\lambda = \min_{\substack{v=0\\ \text{on}\,\partial D}} \frac{(\mathscr{D}(v))^{p+1}}{\int_D v^{2p+2}\, dx}, \tag{3.4}$$

and it is easy to see that the last term on the right in (3.3) gives just the eigenvalue for a disk of area $\bar{A}$, we find the following isoperimetric inequality: *For a given area $\bar{A}$ of D and any $p = 0, 1, 2, \cdots$ the disk yields the lowest eigenvalue λ in the problem* (3.1).

Remarks. (a) For $p = 0$ this is the so-called Faber–Krahn inequality [18].

(b) In the above proof the essential part is that there is a radially symmetric solution for the given "nonlinearity" $f(u)$ and eigenvalue λ for a disk. Clearly, the same statement as above is true for any such $f(u)$. Assuming the existence of a radial symmetric solution, Bandle [3] has given a different result for which the above is a special case.

(c) Our theorem is also closely related to a result of Bandle and Hersch [5] for $f(u)$ increasing and $f(0)>0$. Their result is that the critical value λ^* (see § 1) is a minimum for a disk if the area $\bar{A}$ of D is given.

A possible application of our result is the following. From (3.4) it follows that for any function $v \in C^1(D) \cap L^{2p+2}(D)$ which vanishes on ∂D, we have

$$\int_D v^{2p+2}\, dx \leqq \frac{1}{\lambda} (\mathscr{D}(v))^{p+1}. \tag{3.5}$$

Since we can use any lower bound for λ, we can write e.g. for $p = 1$

$$\int_D v^4\, dx \leqq \frac{1}{\lambda_D} (\mathscr{D}(v))^2, \tag{3.6}$$

where λ_D ($\doteq 14\pi^2/\bar{A}$) is the eigenvalue for a disk of area $\bar{A}$. Since the equality sign holds in (3.6) when D is a disk, we have an optimal Sobolev-type inequality. Such inequalities play an important role in uniqueness and stability criteria for the Navier–Stokes equation.

Let us return now to the inequality (2.6) and more general functions $f(u)$. A multiplication by $4\pi A f(u)$ and integration gives, with $F(u) = \int_0^u f(t)\, dt$,

$$\int_D F(u)\, dx \leqq \frac{\lambda}{8\pi}\left(\int_D f(u)\, dx\right)^2. \tag{3.7}$$

This inequality has been obtained in [22] in a slightly different manner. As remarked in [22], a number of known isoperimetric inequalities are contained in (3.7) as special cases. For example, if $\lambda f(u) \equiv 2$, then (3.7) becomes

$$S \leqq \frac{\bar{A}^2}{2\pi}, \quad S \equiv \mathscr{D}(u) = \text{torsional rigidity of } D,$$

which is the St. Venant inequality (see [15]), while for $f(u)=u$, (3.7) reads

$$\int_D u^2\,dx \leqq \frac{\lambda}{4\pi}\left(\int_D u\,dx\right)^2. \tag{3.8}$$

This converse "Schwarz inequality" for the first eigenfunction in the membrane problem was proven by Payne and Rayner [19]. Actually, their proof was more complicated. For the solution of (3.1) we can write (3.7) after a short calculation as

$$\left(\oint_{\partial D} \frac{\partial u}{\partial n}\,ds\right)^2 \geqq \frac{4\pi}{p+1}. \tag{3.9}$$

In the torsion problem (i.e. $\lambda f(u) \equiv 2$) we can write (2.6) as

$$\frac{du}{dA} + \frac{1}{2\pi} \geqq 0, \quad \text{a.e. in } (0, \bar{A}). \tag{3.10}$$

Multiplying (3.10) by an arbitrary positive function $g(u)$, and integrating gives

$$G(u_{\max}) \leqq \frac{1}{2\pi}\int_D g(u)\,dx, \quad \text{where } G(u) = \int_0^u g(v)\,dv. \tag{3.11}$$

For $g(u)=u$ this gives $u_{\max}^2 \leqq S/(2\pi)$, as observed by Payne [18]. We will make use of (3.11) in the next section. Let us conclude this section by showing an example of a useful nonisoperimetric inequality which also follows from (2.6). If we multiply (2.6) by $dE/dA = f(u(A))$ and integrate again we first are led to

$$F(u_{\max}) \leqq \frac{\lambda}{8\pi}\left\{\frac{1}{\bar{A}}E^2(\bar{A}) + \int_0^{\bar{A}} \left(\frac{E}{A}\right)^2 dA\right\}. \tag{3.12}$$

In the last term on the right we use Hardy's inequality [8], which in this case gives

$$\int_0^{\bar{A}} \left(\frac{E}{A}\right)^2 dA < 4\int_0^{\bar{A}} \left(\frac{dE}{dA}\right)^2 dA. \tag{3.13}$$

Thus, we finally have

$$F(u_{\max}) \leqq \frac{\lambda}{8\pi}\left\{\frac{1}{\bar{A}}\left(\int_S f(u)\,dx\right)^2 + 4\int_D f^2(u)\,dx\right\}. \tag{3.14}$$

A useful application of (3.14) is the following. For $f(u) = u =$ first eigenfunction of the membrane problem, an important physical quantity is the so-called "average-to-peak ratio" ε defined as (see [21])

$$\varepsilon = \int_D u\,dx/(u_{\max}\bar{A}). \tag{3.15}$$

From (3.14) and (3.8) we find then

$$\varepsilon \geqq \left[\frac{\lambda\bar{A}}{4\pi}\left(1+\frac{\lambda\bar{A}}{\pi}\right)\right]^{-1/2}. \tag{3.16}$$

Upper bounds for ε were given in [21], [24]. For a disk (3.16) gives $\varepsilon \geqq 0.32$, while actually $\varepsilon \doteq 0.43$ in this case, whereas for a square (3.16) yields $\varepsilon \geqq 0.30$ and $\varepsilon \doteq 0.40$.

4. Additional inequalities in problem (3.1). In this section we give other inequalities for problem (3.1) which do not follow from (2.6). We first consider a simply connected "starshaped" domain D, i.e.

$$h(s) \equiv (x, n) > 0, \tag{4.1}$$

for some choice of the origin in D, where x also denotes the radius vector of a point x on ∂D, and n is a unit outward normal vector to ∂D at x. A simple application of Green's identity yields for any solution of (1.1) in N dimensions the identity

$$\frac{1}{2}\oint_{\partial D} h(s)\left(\frac{\partial u}{\partial n}\right)^2 ds + \frac{N-2}{2}\mathscr{D}(u) = \lambda N \int_D F(u)\, dx. \tag{4.2}$$

This Rellich-type identity has been used by many authors. We define now

$$B = \oint_{\partial D} h^{-1}(s)\, ds. \tag{4.3}$$

It follows then from Schwarz's inequality and (4.2) that for the solution of (3.1) we have (for $N = 2$)

$$\left(\oint_{\partial D} \frac{\partial u}{\partial n}\, ds\right)^2 \leqq \frac{2B}{p+1} \quad \text{(equality holds for } D \text{ a disk).} \tag{4.4}$$

Remarks. (a) For $p = 0$ (4.4) can be written as

$$\left(\int_D u\, dx\right)^2 \leqq \frac{2B}{\lambda}\int_D u^2\, dx, \tag{4.5}$$

which is sharper than Schwarz's inequality and is not hard to check.

(b) From (4.6) and (3.9) we get the well-known inequality $B \geqq 2\pi$, with equality for a circle.

Next we mention how conformal mapping can be used to get an isoperimetric upper bound for λ in (3.1). Let $w(z) = z - a + \sum_{k=2}^{\infty} c_k(z-a)^k$ be the analytic function which maps the domain D in the z-plane onto the disk of radius R in the w-plane. Let $\tilde{u}(r) = \tilde{u}(|w|)$ be the solution of (3.1) for the disk. Define the "transplanted" function $v(z)$ by $v(z) = \tilde{u}(w(z))$. We then use $v(z)$ as a trial function in the variational characterization (3.4) of λ. As was pointed out in [9] we have, if $\tilde{u} = \tilde{u}(r)$,

$$\int_D v^{2p+2}\, dx \geqq 2\pi \int_0^R \tilde{u}^{2p+2}(r)\, dr. \tag{4.6}$$

Hence by the conformal invariance of the Dirichlet integral and (4.6) we have

$$\lambda_D \leqq \frac{(D(v))^{p+1}}{\int_D v^{2p+2}\, dx} \leqq \frac{1}{2\pi\int_0^R \tilde{u}^{2p+2}(r) r\, dr} = \lambda_C. \tag{4.7}$$

Here λ_D denotes the eigenvalue λ of (3.1) for the given domain D and, correspondingly, λ_C for a disk. We can of course choose the point $a \in D$ such that the radius R is maximal $(= \dot{R})$ and (4.7) can be stated as: *For given maximum conformal radius $\dot{R}$ of D, the eigenvalue λ in* (3.1) *is a maximum for the circle.* For $p = 0$ this is a well-known theorem of Pólya and Szegö [30].

As an application of (3.11), let us finally mention another upper bound for λ. We take as a trial function v in (3.4) the torsion function t, i.e. the solution of (1.1) for $\lambda f(u) \equiv 2$. If we use then (3.11) with $g(t) = t^{2p+2}$ we find

$$\lambda < \frac{S^{p+1}}{t_{\max}^{2p+3}} \frac{2p+3}{2\pi}. \tag{4.8}$$

Various bounds for S and $t_{\max}$ can be found in [18]. For $p = 1$ and D a disk of radius R, (4.8) gives $\lambda < 20\pi/R^2$.

5. Concluding remarks. (a) We first mention that one can derive in the same way an analogue to (2.6) in N-dimensions, where one has to use the corresponding isoperimetric inequality between N-volume and surface area. However some of the integrations get considerably more complicated, as shown in [19].

(b) In the "inhomogeneous" case

$$\begin{aligned} &\Delta u + \lambda \rho(x) f(u) = 0 \quad && \text{in } D \subset R^2 \\ &u = 0 && \text{on } \partial D, \end{aligned} \tag{5.1}$$

one could derive another analogue of (2.6) using the ideas of C. Bandle (see e.g. [2]) to get then

$$\frac{du}{dm} + \frac{\lambda E}{(\pi - km)m} \geqq 0, \quad \text{for } m \in (0, M), \tag{5.2}$$

where $m = \int_{D(t)} \rho\, dx$, $M = \int_D \rho\, dx$, $D(t)$ as before and $E(m) = \int_0^m u(y)\, dy$. Here, one has to assume that $\Delta(\log \rho) + 2k\rho \geqq 0$ in D and can then use an inequality of Alexandrov instead of the classical isoperimetric inequality. For details we refer the reader again to [2].

(c) There are possible extensions of the techniques used here to the case that we have boundary conditions of the third kind for u in (1.1). These extensions as indicated in [28] are not yet satisfactory.

REFERENCES

[1] H. AMANN, *On the existence of positive solutions of nonlinear elliptic boundary-value problems*, Indiana Univ. Math. J., 21 (1971), pp. 215–246.

[2] C. BANDLE, *Konstruktion isoperimetrischer Ungleichungen der mathematischen Physik aus solchen der Geometrie*, Comm. Math. Helv., 46 2 (1971), pp. 182–213.

[3] ———, *The Rayleigh–Faber–Krahn theorem for the characteristic values associated with a class of nonlinear boundary value problems*, this Journal, 4 (1973), pp. 8–14.

[4] ———, *Existence theorems, some qualitative results and a priori bounds for a class of nonlinear Dirichlet problems*, Arch. Rational Mech. Anal., 58 (1976), pp. 219–238.

[5] C. BANDLE AND J. HERSCH, *Problèmes de Dirichlet non linéaires: Une condition suffisante isopérimétrique pour l'existence d'une solution*, C. R. Acad. Sci. Paris, 280(1975), pp. 1057–1060.

[6] S. BERNSTEIN, *Sur la nature analytique des solutions aux dérivées partielles du second ordre*, Math. Ann., 59 (1904), pp. 20–76.

[7] P. CROOKE, *On two inequalities of Sobolev type*, Applicable Anal., 3 (1974), pp. 345–358.

[8] ———, On an improved Sobolev constant, to appear.

[9] I. M. GEL'FAND, *Some problems in the theory of quasilinear equations*, Amer. Math. Soc. Transl. (2), 29 (1963), pp. 295–381.

[10] G. H. HARDY, *Note on a theorem of Hilbert*, Math. Z., 6 (1920), pp. 314–317.

[11] J. HERSCH, *On symmetric membranes and conformal radius: Some complements to Pólya's and Szegö's inequalities*, Arch. Rational Mech. Anal., 20 (1965), pp. 378–390.

[12] S. I. HUDJAEV, *Boundary-value problems for certain quasilinear elliptic equations*, Soviet Math. Dokl., 5 (1964), pp. 188–192.

[13] D. D. JOSEPH AND T. S. LUNDGREN, *Quasilinear Dirichlet problems driven by positive sources*, Arch. Rational Mech. Anal., 49 (1973), pp. 241–269.

[14] H. B. KELLER AND D. S. COHEN, *Some positone problems suggested by nonlinear heat generations*, J. Math. Mech., 16 (1967), pp. 1361–1376.

[15] T. LAETSCH, *A note on a paper of Keller and Cohen*, Ibid., 18 (1969), pp. 1095–1099.

[16] N. LEVINSON, *Positive eigenfunctions for* $\Delta u+\lambda f(u)=0$, Arch. Rational Mech. Anal., 11 (1962), pp. 1065–1072.

[17] L. NIRENBERG, *On nonlinear partial differential equations and Hölder continuity*, Comm. Pure Appl. Math., 6(1953), pp. 103–156.

[18] L. E. PAYNE, *Isoperimetric inequalities and their applications*, SIAM Review, 9 (1967), pp. 453–488.

[19] L. E. PAYNE AND M. E. RAYNER, *An isoperimetric inequality for the first eigenfunction in the fixed membrane problem*, Z. Angew. Math. Phys., 23 (1972), pp. 13–15.

[20] ———, *Some isoperimetric norm bounds for solutions of the Helmholtz equation*, Ibid., 24 (1973), pp. 105–110.

[21] L. E. PAYNE AND I. STAKGOLD, *On the mean value of the fundamental mode in the fixed membrane problem*, Applicable Anal., 3 (1973), pp. 295–306.

[22] L. E. PAYNE, R. P. SPERB AND I. STAKGOLD, *On Hopf-type maximum principles for convex domains*, Nonlinear Anal., 5 (1977), pp. 547–559.

[23] S. I. POHOZAEV, *Eigenfunctions of the equation* $\Delta u+\lambda f(u)=0$, Soviet Math. Dokl., 6 (1965), pp. 1408–1411.

[24] P. W. SCHAEFER AND R. P. SPERB, *Maximum principles for some functionals associated with the solution of elliptic boundary value problems*, Arch. Rational Mech. Anal., 61 (1976), pp. 65–76.

[25] L. F. SHAMPINE, *Some nonlinear eigenvalue problems*, J. Math. Mech., 17 (1968), pp. 1065–1072.

[26] R. P. SPERB, *Extensions of two theorems of Payne to some nonlinear Dirichlet problems*, Z. Angew. Math. Phys., 26 (1975), pp. 721–726.

[27] ———, *Maximum Principles and Nonlinear Elliptic Problems*, to appear.

[28] I. STAKGOLD, *Global Estimates for Nonlinear Reaction and Diffusion*, Springer Lecture Notes, Springer-Verlag, Berlin, vol. 415, pp. 252–266.

[29] R. E. BELLMAN., *Stability Theory of Differential Equations*, McGraw-Hill, New York, 1953.

[30] G. PÓLYA AND G. SZEGÖ, *Isoperimetric Inequalities in Mathematical Physics*, Annals of Mathematics Studies 27, Princeton University Press, Princeton, NJ, 1951.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0012 $01.00/0

ON THE KAKEYA-ENESTRÖM THEOREM AND GEGENBAUER POLYNOMIAL SUMS*

STEPHAN RUSCHEWEYH†

Abstract. An extension of the classical Kakeya–Eneström theorem is given. As an application we show that for $\lambda \geqq \frac{1}{2}$, $-1<x<1$ and arbitrary nonincreasing sequences $a_k>0$, $k=0, 1, \cdots, n$, we have

$$\sum_{k=0}^{n} a_k \frac{C_k^{(\lambda)}(x)}{C_k^{(\lambda)}(1)} z^k \neq 0, \qquad |z| \leqq 1,$$

where $C_k^{(\lambda)}$ are the Gegenbauer or ultraspherical polynomials. This extends an old result due to G. Szegö and settles two recent conjectures of R. Askey and J. Bustoz. Other related results are obtained as well.

1. Introduction. Let $a_k \in \mathbb{R}$, $k=0, 1, \cdots, n$, be such that

$$a_0 \geqq a_1 \geqq \cdots \geqq a_n > 0.$$

The Kakeya–Eneström theorem (see [4, p. 136]) states that the polynomial $a_0+a_1z+\cdots+a_nz^n$ has no zeros inside the unit disk $\Delta=\{z \| z|<1\}$. In the present paper we shall give a far-reaching extension of this classical result.

Let $\mathscr{A}$ denote the class of functions $f(z)$ analytic in Δ with $f(0)=0, f'(0)=1$. $f \in \mathscr{A}$ is called *starlike of order* β, $\beta \leqq 1$, if and only if

$$\operatorname{Re} \frac{zf'(z)}{f(z)} \geqq \beta, \qquad z \in \Delta,$$

and $\mathscr{S}_\beta$ is the collection of these functions. Important members of $\mathscr{S}_{1/2}$ are the functions

$$e_x(z) = \frac{z}{1-xz}, \qquad |x|=1,$$

which represent the extreme points of the closed convex hull of $\mathscr{S}_{1/2}$.

THEOREM 1. *Let $n \in \mathbb{N}$ and $f(z)=z\sum_{k=0}^{\infty} b_k z^k \in \mathscr{S}_{1/2}$. Then there exists a number $\rho=\rho(n,f) \geqq 1$ such that for every sequence $a_k \in \mathbb{R}$, $k=0, 1, \cdots, n$, with*

$$1=a_0 \geqq a_1 \geqq \cdots \geqq a_n \geqq 0,$$

we have

$$P(z)=\sum_{k=0}^{n} a_k b_k z^k \neq 0, \qquad |z|<\rho. \tag{1}$$

For $f \neq e_x$, $|x|=1$, we have $\rho(n,f)>1$.

Remarks. 1) Note that the Kakeya–Eneström Theorem corresponds to the cases $f=e_x$, $|x|=1$.

2) It is obvious from Theorem 1 that if

$$\min_{k=1,\cdots,n} \frac{a_{k-1}}{a_k} = \tilde{\rho} > 1,$$

we actually have $P(z) \neq 0$ for $|z|<\rho\tilde{\rho}$.

* Received by the editors August 30, 1976, and in final revised form December 2, 1976.
† Department of Mathematics, University of Würzburg, 8700 Würzburg, West Germany.

3) Under the assumptions of Theorem 1, one can prove the following stronger result: *There exists a number* $\rho'(n,f)\geqq 1$ *such that* $g(z)=\int_0^z P(\zeta)\,d\zeta$ *is univalent in* $|z|<\rho'$. *For* $f\neq e_x$, $|x|=1$, *we have* $\rho'(n,f)>1$.

Example. The choice

$$a_k=\binom{n-k+\alpha}{n-k}\bigg/\binom{n+\alpha}{n}, \qquad \alpha>0,$$

leads to the conclusion that the nth Cesaro mean of order α of a function $f\in\mathcal{S}_{1/2}$ is nonvanishing for $0<|z|<1+\alpha/n$.

We mention a few properties of $\mathcal{S}_{1/2}$.

(i) If

$$z\sum_{k=0}^{\infty} b_k z^k\in\mathcal{S}_{1/2}, \qquad z\sum_{k=0}^{\infty} c_k z^k\in\mathcal{S}_{1/2}$$

then

$$z\sum_{k=0}^{\infty} b_k c_k z^k\in\mathcal{S}_{1/2},$$

(see [5, Thm. 3.1]).

(ii) If $f(z)=z\sum_{k=0}^{\infty} b_k z^k$ is prestarlike of order $\beta\leqq 1/2$, i.e.

$$z\sum_{k=0}^{\infty}\binom{2\beta-2}{k} b_k z^k\in\mathcal{S}_\beta,$$

then $f\in\mathcal{S}_{1/2}$. In particular every function $f\in\mathcal{A}$ which maps Δ univalently onto a convex domain is prestarlike of order zero and hence in $\mathcal{S}_{1/2}$ (see [7, Thm. 10], [6]).

(iii) If $f(z)=z\sum_{k=0}^{\infty} b_k z^k$ is prestarlike of order $\beta\leqq 1-n/2$, $n\in\mathbb{N}$, then $z\sum_{k=0}^{\infty} b_{nk} z^k\in\mathcal{S}_{1/2}$. This is an obvious consequence of the results in [6].

The functions

$$\frac{z}{(1-2xz+z^2)^\lambda}=z\sum_{k=0}^{\infty} C_k^{(\lambda)}(x) z^k, \qquad -1\leqq x\leqq 1, \quad \lambda\geqq 0,$$

where $C_k^{(\lambda)}$ are the Gegenbauer polynomials, are obviously in $\mathcal{S}_{1-\lambda}$. Since

$$C_k^{(\lambda)}(1)=(-1)^k\binom{-2\lambda}{k},$$

it follows from (ii) that the functions

$$z\sum_{k=0}^{\infty}\frac{C_k^{(\lambda)}(x)}{C_k^{(\lambda)}(1)} z^k \tag{2}$$

are prestarlike of order $1-\lambda$ and hence in $\mathcal{S}_{1/2}$ for $\lambda\geqq\frac{1}{2}$.

This observation together with the property mentioned in (iii) gives the following Corollary to Theorem 1.

COROLLARY. *Let* $\lambda\geqq m/2$, $m\in\mathbb{N}$. *Let* $a_k\in\mathbb{R}$, $k=0,1,\cdots,n$, *satisfy*

$$1=a_0\geqq a_1\geqq\cdots\geqq a_n\geqq 0.$$

Then for $-1<x<1$ *we have*

$$\sum_{k=0}^{n} a_k\frac{C_{km}^{(\lambda)}(x)}{C_{km}^{(\lambda)}(1)} z^k\neq 0, \qquad |z|\leqq 1. \tag{3}$$

Remarks. 4) The case $\lambda=\frac{1}{2}$, $a_k=1$ for $k=0,1,\cdots,n$ is G. Szegö's well known result [8]. The cases $\lambda\geqq\frac{1}{2}$, $m=1$, and

$$a_k=\binom{n-k+\alpha}{n-k}\Big/\binom{n+\alpha}{n}$$

have been conjectured by R. Askey [1] ($\alpha=0$) and J. Bustoz [2] ($\alpha>0$) respectively.

5) It should be noted that (3) contains a positivity result: For $\lambda\geqq m/2$, $m\in\mathbb{N}$, we have

$$\sum_{k=0}^{n}\frac{C_{km}^{(\lambda)}(x)}{C_{km}^{(\lambda)}(1)}>0,\qquad -1<x<1.$$

2. Proofs. Let $\mathcal{P}_\beta$, $\beta\leqq 1$, denote the class of functions $f(z)$ analytic in Δ with $f(0)=1$ and $\operatorname{Re} f(z)\geqq\beta$, $z\in\Delta$. We shall require the following results.

LEMMA 1. *Let $f(z)=\sum_{k=0}^{\infty}a_kz^k\in\mathcal{P}_{1/2}$. Then $|a_k|\leqq 1$, $k\in\mathbb{N}$. If equality takes place for at least one k_0 then there exists $\theta_0\in\mathbb{R}$ such that for all $m\in\mathbb{N}$, $k=0,1,\cdots,k_0-1$,*

$$a_{k+mk_0}=e^{i\theta_0 mk_0}a_k. \tag{4}$$

LEMMA 2. *For $f\in\mathcal{P}_0$ we have*

$$|f(z)|\geqq\frac{1-|z|}{1+|z|},\qquad z\in\Delta. \tag{5}$$

LEMMA 3. *$f\in\mathcal{A}$ is in $\mathcal{S}_{1/2}$ if and only if for every $z_0\in\Delta$*

$$h_{z_0}(z)=\frac{z_0}{f(z_0)}\,\frac{f(z)-f(z_0)}{z-z_0}\in\mathcal{P}_{1/2}. \tag{6}$$

LEMMA 4. *Let $f(z)=z\sum_{k=0}^{\infty}b_kz^k\in\mathcal{S}_{1/2}$, $f\neq e_x$ for $|x|=1$. Then $\lim_{k\to\infty}b_k=0$.*

Lemmata 1, 2 are well known. Lemma 3 is in [5, Thm. 1.5]. Lemma 4 follows from the fact that $\mathcal{S}_{1/2}\backslash\{e_x\,|\,|x|=1\}$ is contained in the Hardy space $\mathcal{H}^1$ (see [3, Thm. 5]).

Proof of Theorem 1. Define

$$\pi_j(z)=\sum_{k=0}^{j}b_kz^k,\qquad j=0,1,\cdots,$$

such that $P(z)$ has a uniquely determined representation

$$P(z)=\sum_{j=0}^{n}\alpha_j\pi_j(z),\qquad \alpha_j\geqq 0,\qquad \sum_{j=0}^{n}\alpha_j=1. \tag{7}$$

This shows that the polynomials under consideration form exactly the closed convex hull of the set $\{\pi_j(z)\,|\,j=0,1,\cdots,n\}$. From (6) we obtain $h_{z_0}(z)=\sum_{j=0}^{\infty}A_j(z_0)z^j\in\mathcal{P}_{1/2}$ with

$$A_j(z_0)=\begin{cases}1, & j=0\\ \dfrac{1}{z_0^j}\left(1-\dfrac{z_0\pi_{j-1}(z_0)}{f(z_0)}\right), & j\in\mathbb{N}.\end{cases}$$

Lemma 1 shows $|A_j(z_0)|\leqq 1$ and hence

$$\operatorname{Re}\frac{z_0\pi_{j-1}(z_0)}{f(z_0)}>0,\qquad j\in\mathbb{N},\quad z_0\in\Delta.$$

This gives

(8) $$\operatorname{Re}\frac{z_0P(z_0)}{f(z_0)}>0, \qquad z_0\in\Delta,$$

and therefore $P(z)\neq 0$, $\quad z\in\Delta$.

To complete the proof it remains to show $P(z)\neq 0$ on $|z|=1$ whenever $f\neq e_x$, $|x|=1$. The existence of the number $\rho(n,f)>1$ then follows from the fact that the considered set of polynomials is closed.

Let $P(\zeta)=0$ for a certain $\zeta\in\partial\Delta$. There is no loss of generality if we assume $\zeta=1$. Since $\mathscr{P}_{1/2}$ is a normal family we can choose a sequence $z_k\in(0,1)$, $z_k\to 1$, such that

(i) $f(z_k)$ tends to a finite or infinite limit α, and

(ii) $h_{z_k}(z)$ is compact convergent to $h(z)\in\mathscr{P}_{1/2}$.

We first exclude the case $\alpha=\infty$. In fact, $\alpha=\infty$ would imply

$$\lim_{k\to\infty}\frac{z_kP(z_k)}{(1-z_k)f(z_k)}=0,$$

which contradicts the inequality

$$\left|\frac{z_kP(z_k)}{(1-z_k)f(z_k)}\right|\geqq\frac{1}{1+|z_k|}$$

following from (8) and Lemma 2. Hence $\alpha\neq\infty$, and we obtain

$$h(z)=\frac{1}{\alpha}\,\frac{f(z)-\alpha}{z-1}=1+\sum_{j=1}^{\infty}\left(1-\frac{\pi_{j-1}(1)}{\alpha}\right)z^j$$
$$=\sum_{j=0}^{\infty}A_jz^j\in\mathscr{P}_{1/2}.$$

By Lemma 1 we conclude

(9) $$\left|\frac{1-\pi_{j-1}(1)}{\alpha}\right|\leqq 1, \qquad j\in\mathbb{N};$$

but on the other hand

$$\sum_{j=1}^{n+1}\alpha_{j-1}\left(1-\frac{\pi_{j-1}(1)}{\alpha}\right)=\sum_{j=1}^{n+1}\alpha_{j-1}-\frac{P(1)}{\alpha}=1.$$

This shows that equality must occur in (9) for at least one index. Since

$$f(z)=\alpha\sum_{j=1}^{\infty}(A_{j-1}-A_j)z^j,$$

we deduce from (4) that the sequence

$$|A_{j-1}-A_j|, \qquad j\in\mathbb{N},$$

is periodic, which, however, contradicts Lemma 4.

3. Finally we wish to point out an easy consequence of Lemma 3: For $f(z)=z\sum_{k=0}^{\infty}b_kz^k\in\mathscr{S}_{1/2}$ we have

$$\left|\sum_{k=n}^{\infty}b_kz^k\right|<\left|\sum_{k=0}^{\infty}b_kz^k\right|, \qquad z\in\Delta, \quad n\in\mathbb{N}.$$

This result, for

$$f_0(z)=z\sum_{k=0}^{\infty}\frac{C_k^{(\lambda)}(x)}{C_k^{(\lambda)}(1)}z^k,\qquad \lambda\geqq\tfrac{1}{2},\quad -1\leqq x\leqq 1,$$

has previously been obtained by R. Askey [1] in the case $z=r\in(-1,1)$. Since $f_0\in\mathcal{S}_{1/2}$ Askey's result now extends to $z\in\Delta$.

REFERENCES

[1] R. Askey, *Positive Jacobi polynomials*, Tôhoku Math. J., 24 (1972), pp. 109–119.

[2] J. Bustoz, *Jacobi polynomial sums and univalent Cesàro means*, Proc. Amer. Math. Soc., 50 (1975), pp. 259–264.

[3] P. J. Eenigenburg and F. R. Keogh, *The Hardy class of some univalent functions and their derivatives*, Michigan Math. J. 17 (1970), pp. 335–346.

[4] M. Marden, *Geometry of Polynomials*, Mathematical Surveys 3, American Mathematical Society, Providence, RI, 1966.

[5] St. Ruscheweyh and T. Sheil-Small, *Hadamard products of schlicht functions and the Pólya–Schoenberg conjecture*, Comment. Math. Helv., 48 (1973), pp. 119–135.

[6] St. Ruscheweyh, *Linear operators between classes of prestarlike functions*, Ibid., to appear.

[7] T. J. Suffridge, *Starlike functions as limits of polynomials*, Advances in Complex Function Theory, Lecture Notes in Mathematics 505, Springer-Verlag, Berlin-Heidelberg-New York, 1976, pp. 164–202.

[8] G. Szegö, *Zur Theorie der Legendreschen Polynome*, Jber. Deutsch. Math.-Verein., 40 (1931), pp. 163–166.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0013 $01.00/0

ASYMPTOTIC BEHAVIOR OF SOLUTIONS OF A CLASS OF NONLINEAR DIFFERENTIAL EQUATIONS IN BANACH SPACE*

TAKAO NAMBU†

Abstract. A semi-linear ordinary differential equation in a Banach space is considered. The coefficient operator $A(t)$ has the domain $D(A)$ which is independent of t and not necessarily dense. It is shown that the evolution operator $U(t, s)$ corresponding to the unperturbed linear equation has an integrable singularity at $t=s$ and is strongly continuously differentiable in s on $D(A)$. Such examples have been obtained by W. von Wahl [13] and H. Kielhöfer [4], [5], and are also obtained in this paper. The nonlinear term satisfies either a uniform or local Lipschitz's condition with respect to the unknown solution. The principal tools are the semigroup theory and integral inequalities. Several results on the asymptotic behavior of the solution of the semi-linear equation are obtained. These results are applicable to the problem of the stability of semi-linear parabolic initial boundary value problems within the framework of the C^α-theory.

1. Introduction. The purpose of this paper is to study the problems of existence, uniqueness, and asymptotic behavior of solutions of the following semi-linear evolution equation in a Banach space E:

$$\begin{aligned} &\frac{d}{dt}x(t)=A(t)x(t)+f(t,x(t)), \qquad 0<t<+\infty, \\ &x(0)=x_0. \end{aligned} \tag{1.1}$$

$x(t)$ is said to be a solution of (1.1), if

$$x(\cdot)\in C([0,\infty);E)\cap C^1((0,\infty);E)$$

and (1.1) is satisfied.

We assume that $A(t)$, $t\geqq 0$, in (1.1) is a closed linear operator in E and that the domain of $A(t)$, which will be denoted by $D(A)$, does not depend on t and is not necessarily dense in E. Throughout this paper, it is assumed that the resolvent $(A(t)-\mu I)^{-1}$ exists and satisfies the estimate

$$\|(A(t)-\mu I)^{-1}\|\leqq \frac{M_0}{(1+|\mathrm{Im}\,\mu|)^\beta}, \tag{1.2}$$

for $\mathrm{Re}\,\mu\geqq -\lambda$ and $t\geqq 0$, where $0<\beta<1, \lambda>0$, and M_0 are independent of t. Furthermore it is assumed that the estimate

$$\|(A(t)-A(s))A^{-1}(0)\|\leqq L_1(t-s)^\rho, \qquad 0\leqq s\leqq t\leqq T, \tag{1.3}$$

holds for each $T>0$, where $0<\rho\leqq 1$, and L_1 is independent of t and s. (In §§ 4, 5, and 6, we impose a stronger assumption on L_1.)

W. von Wahl [13] and H. Kielhöfer [4], [5] studied the parabolic initial boundary value problems, and obtained the similar estimate to (1.2). The estimate (1.2) is also obtained for some type of parabolic systems (see § 7).

For fixed q, $0<q<1$, let Γ_q be the curve

$$\Gamma_q=\left\{\mu=\sigma+i\tau;\sigma=-\lambda-\frac{q}{M_0}(1+|\tau|)^\beta,\ -\infty<\tau<+\infty\right\}.$$

* Received by the editors June 4, 1976, and in revised form December 31, 1976.
† Faculty of Engineering Science, Osaka University, Toyonaka, Osaka, Japan.

It is easy to see that the resolvent of $A(t)$ exists in a region situated to the right of the curve Γ_q and satisfies the estimate (1.2) with $M_0(1-q)^{-1}$ instead of M_0 [6].

From (1.2), for each $t \geqq 0$ the *weakened Cauchy problem* of the linear equation

$$\frac{d}{dh} x(h) = A(t)x(h), \quad h > 0, \qquad x(0) = x_0, \tag{1.4}$$

with a constant coefficient $A(t)$ is well posed on the set $D(A)$ [6], and the solution $x(h)$ of (1.4) is represented as $e^{hA(t)}x_0$ for $h > 0$, where $e^{hA(t)}$, $h > 0$, is the semigroup of bounded linear operators given by

$$e^{hA(t)} = -\frac{1}{2\pi i}\int_{\Gamma_q} e^{\mu h}(A(t)-\mu I)^{-1}\, d\mu, \qquad h > 0. \tag{1.5}$$

It follows from (1.5) that, for each $t \geqq 0$, $e^{hA(t)}$ is infinitely continuously differentiable in norm in $h > 0$, and satisfies the estimates

$$\begin{gathered} \|A^n(t)\, e^{hA(t)}\| \leqq M_{n+1} \cdot e^{-\lambda h} \cdot h^{1-(n+1)/\beta}, \\ n = 0, 1, \cdots, \quad h > 0, \quad t \geqq 0, \end{gathered} \tag{1.6}$$

where M_{n+1} are independent of t and h [6].

In the next section, the following unperturbed linear equation will be considered:

$$\frac{\partial}{\partial t} x(t, s) = A(t)x(t, s), \qquad t > s, \quad x(s, s) = x_0. \tag{1.7}$$

S. G. Krein [6] and E. T. Poulsen [8] constructed the evolution operator $U(t, s)$ of (1.7) under stronger conditions than in this paper. We will construct $U(t, s)$, following H. Tanabe [10], [12] and it will be shown that $U(t, s)$ is unique and has an integrable singularity at $t = s$.

For the nonlinear term $f(t, x)$, it is assumed that the following condition (i) or (ii) is satisfied:

(i) $f(t, x)$ is continuous on $[0, \infty) \times E$, and the estimate

$$\|f(t, x) - f(t, y)\| \leqq K\|x - y\| \tag{1.8}$$

holds, where K is a constant independent of t, x, and y.

(ii) $f(t, x)$ is continuous on $[0, \infty) \times E$. For each $c > 0$, there exists a constant $k(c) > 0$ such that the estimate

$$\|f(t, x) - f(t, y)\| \leqq k(c)\|x - y\| \tag{1.9}$$

holds for t, x, and y satisfying $t \geqq 0$, $\|x\| \leqq c$, $\|y\| \leqq c$.

As quoted above, W. von Wahl [13] and H. Kielhöfer [4], [5] considered the local solvability of the semi-linear evolution equation (1.1) corresponding to the parabolic initial boundary value problem within the framework of the L^p- and C^α-theory under the similar conditions to (1.2) and weaker conditions for $f(t, x)$. In applications, especially when treating the problems of the stability of nonlinear partial differential equations, it is important to consider them within the framework of the C^α-theory as well as the L^p-theory. Because the convergence in L^p does not guarantee the almost everywhere convergence. Our main purpose is to obtain an estimate for the asymptotic behavior of the solution of (1.1) and its derivative under (1.2), (1.3), and either (1.8) or (1.9). Some examples will be worked out.

2. Construction of evolution operator. In this section, we construct the evolution operator $U(t, s)$ of (1.7). From (1.3), it follows that the bounded operator $A(0)A^{-1}(t)$ is continuous in norm for $t \geqq 0$. From (1.2) and (1.3), $A^k(0)\, e^{hA(t)}$ $(k = 0, 1)$ is continuous in norm for $t \geqq 0$ and $h > 0$ [6]. Therefore it follows that the bounded operator

$$R_1(t, s) = (A(t) - A(s))\, e^{(t-s)A(s)}$$

is continuous in norm for $t > s$ and the estimate

$$\|R_1(t, s)\| \leqq c_1(t-s)^{\rho+1-2/\beta}, \qquad 0 \leqq s < t \leqq T, \tag{2.1}$$

holds for each $T > 0$, where $c_1 > 0$ is some constant depending on T.

Let us define $U(t, s)$ as follows [10], [12]:

$$U(t, s) = e^{(t-s)A(s)} + \int_s^t e^{(t-\tau)A(\tau)} R(\tau, s)\, d\tau, \tag{2.2}$$

where $R(t, s)$ is the solution of the integral equation

$$R(t, s) = R_1(t, s) + \int_s^t R_1(t, \tau) R(\tau, s)\, d\tau. \tag{2.3}$$

If $\beta > 2/3$ and $\rho > 2(1/\beta - 1)$, then (2.3) can be solved uniquely by the successive approximations:

$$R(t, s) = \sum_{n=1}^{\infty} R_n(t, s),$$

$$R_n(t, s) = \int_s^t R_1(t, \tau) R_{n-1}(\tau, s)\, d\tau.$$

It follows from (2.1) that $R(t, s)$ is continuous in norm for $t > s \geqq 0$ and that the estimate

$$\|R(t, s)\| \leqq c_2(t-s)^{\rho+1-2/\beta}, \qquad 0 \leqq s < t \leqq T, \tag{2.4}$$

holds for each $T > 0$. In the following of this section we show that $U(t, s)$ defined in (2.2) is the evolution operator of (1.7) by posing some additional assumption on ρ and β. Throughout this paper, it is assumed that the following condition is satisfied:

$$\sqrt{2/3} < \beta < 1, \qquad \frac{2}{\beta}\left(\frac{1}{\beta} - \beta\right) < \rho \leqq 1. \tag{2.5}$$

Remark. The condition (2.5) is weakened in Example 7 of § 7.

We start with the following lemma:

LEMMA 2.1. *For any* γ, $0 < \gamma < \beta\rho + 2(\beta - 1)$, *there exist* $c_3 > 0$ *and* $c_4 > 0$ *such that the estimates*

$$\|R(t, s) - R(\tau, s)\| \leqq c_3(t-\tau)^{\gamma}(\tau - s)^{\rho+1-2/\beta-\gamma/\beta}, \tag{2.6}$$

$$\|(R(t, s) - R(\tau, s))A^{-1}(s)\| \leqq c_4(t-\tau)^{\gamma}(\tau-s)^{1-1/\beta} \tag{2.7}$$

hold for $0 \leqq s < \tau < t \leqq T$.

Proof. The proof of (2.6) is the same as [10], [12]. The proof of (2.7) can be carried out in the same way as that of (2.6) with some modifications. Hence we omit it.

LEMMA 2.2. *Let us define the bounded operator*

$$S(t,s)=A(t)e^{(t-s)A(t)}-A(s)e^{(t-s)A(s)}, \qquad t>s\geqq 0. \tag{2.8}$$

Then, $S(t,s)$ is continuous in norm for $t>s\geqq 0$ and the estimate

$$\|S(t,s)\|\leqq c_5(t-s)^{\rho+2-3/\beta} \tag{2.9}$$

holds for $0\leqq s<t\leqq T$.

Proof. The proof can be carried out in the same way as [10], [12]. Hence we omit it.

Let any $\delta>0$ and any ε, $0<\varepsilon<\delta/2$, be given. Consider the function

$$W_\varepsilon(t,s)=\int_s^{t-\varepsilon} e^{(t-\tau)A(\tau)}R(\tau,s)\,d\tau. \tag{2.10}$$

Obviously $W_\varepsilon(t,s)$ converges to

$$\int_s^t e^{(t-\tau)A(\tau)}R(\tau,s)\,d\tau$$

uniformly in $t\in[s+\delta,T]$. $W_\varepsilon(t,s)$ is continuously differentiable in $t\in[s+\delta,T]$ and the following equation holds:

$$\begin{aligned}\frac{\partial}{\partial t}W_\varepsilon(t,s)&=e^{\varepsilon A(t-\varepsilon)}R(t-\varepsilon,s)-\int_s^{t-\varepsilon}S(t,\tau)R(\tau,s)\,d\tau\\ &\quad+\int_s^{t-\varepsilon}A(t)\,e^{(t-\tau)A(t)}\{R(\tau,s)-R(t,s)\}\,d\tau\\ &\quad+\{e^{(t-s)A(t)}-e^{\varepsilon A(t)}\}R(t,s).\end{aligned} \tag{2.11}$$

Let $\varepsilon\downarrow 0$. Then, it follows from (2.4) and (2.9) that the second term of (2.11) converges to

$$-\int_s^t S(t,\tau)R(\tau,s)\,d\tau$$

uniformly in $t\in[s+\delta,T]$. From (2.5), we can choose γ in Lemma 2.1 such that the relation

$$2\left(\frac{1}{\beta}-1\right)<\gamma<\beta\rho+2(\beta-1) \tag{2.12}$$

holds. Thus the third term of (2.11) converges to

$$\int_s^t A(t)\,e^{(t-\tau)A(t)}\{R(\tau,s)-R(t,s)\}\,d\tau$$

uniformly in $t\in[s+\delta,T]$ as $\varepsilon\downarrow 0$. As for the first and the fourth terms of (2.11), we note that the following inequality holds:

$$\begin{aligned}&\|e^{\varepsilon A(t-\varepsilon)}R(t-\varepsilon,s)-e^{\varepsilon A(t)}R(t,s)\|\\ &\leqq\|e^{\varepsilon A(t-\varepsilon)}\{R(t-\varepsilon,s)-R(t,s)\}\|+\|\{e^{\varepsilon A(t-\varepsilon)}-e^{\varepsilon A(t)}\}R(t,s)\|\\ &\leqq c_6\varepsilon^{1-1/\beta+\gamma}\\ &\quad+\left\|-\frac{1}{2\pi i}\int_{\Gamma_q}e^{\varepsilon\mu}\{(A(t-\varepsilon)-\mu I)^{-1}-(A(t)-\mu I)^{-1}\}\,d\mu\right\|\|R(t,s)\|\\ &\leqq c_6\varepsilon^{1-1/\beta+\gamma}+c_7\varepsilon^{\rho+2-2/\beta}, \qquad s+\delta\leqq t\leqq T.\end{aligned}$$

Therefore we find that

$$\int_s^t e^{(t-\tau)A(\tau)}R(\tau, s)\,d\tau$$

is continuously differentiable in norm for t ($>s$) and so is $U(t, s)$. Thus we obtain the following equation:

$$\begin{aligned}\frac{\partial}{\partial t}U(t,s)=A(s)\,e^{(t-s)A(s)}&-\int_s^t S(t,\tau)R(\tau,s)\,d\tau\\ &+\int_s^t A(t)\,e^{(t-\tau)A(t)}\{R(\tau,s)-R(t,s)\}\,d\tau\\ &+e^{(t-s)A(t)}R(t,s).\end{aligned}\tag{2.13}$$

Next we show that the range of $U(t, s)$ is contained in $D(A)$. $U(t, s)$ can be expressed as follows:

$$\begin{aligned}U(t,s)=e^{(t-s)A(s)}&+\int_s^t e^{(t-\tau)A(\tau)}\{R(\tau,s)-R(t,s)\}\,d\tau\\ &+\int_s^t \{e^{(t-\tau)A(\tau)}-e^{(t-\tau)A(t)}\}R(t,s)\,d\tau\\ &+\{e^{(t-s)A(t)}A^{-1}(t)-A^{-1}(t)\}R(t,s).\end{aligned}\tag{2.14}$$

It follows from (2.12) that $\|A(t)e^{(t-\tau)A(\tau)}\{R(\tau,s)-R(t,s)\}\|$ is integrable on the interval (s, t). We note that the relation

$$\begin{aligned}&\|A(t)\{e^{(t-\tau)A(t)}-e^{(t-\tau)A(\tau)}\}\|\\ &=\left\|-\frac{1}{2\pi i}\int_{\Gamma_q} e^{(t-\tau)\mu}A(t)(A(t)-\mu I)^{-1}(A(\tau)-A(t))(A(\tau)-\mu I)^{-1}\,d\mu\right\|\\ &\leqq c_8(t-\tau)^{\rho+2-3/\beta}\end{aligned}$$

holds. Consequently it follows that $U(t,s)E\subset D(A)$ and that the equation

$$\begin{aligned}A(t)U(t,s)=A(t)\,e^{(t-s)A(s)}&+\int_s^t A(t)\,e^{(t-\tau)A(\tau)}\{R(\tau,s)-R(t,s)\}\,d\tau\\ &+\int_s^t A(t)\{e^{(t-\tau)A(\tau)}-e^{(t-\tau)A(t)}\}R(t,s)\,d\tau\\ &+\{e^{(t-s)A(t)}-I\}R(t,s)\end{aligned}\tag{2.15}$$

holds. It follows from (2.13) and (2.15) that

$$\begin{aligned}&\frac{\partial}{\partial t}U(t,s)-A(t)U(t,s)\\ &\qquad=R(t,s)-R_1(t,s)-\int_s^t R_1(t,\tau)R(\tau,s)\,d\tau=0.\end{aligned}$$

Noting the estimate $\|R(t,s)A^{-1}(0)\|\leqq c_9(t-s)^{\rho+1-1/\beta}$ and the fact that $e^{hA(t)}A^{-1}(0)$ converges to $A^{-1}(0)$ in norm as $h\downarrow 0$ uniformly in t [6], we find that

$$U(t,s)A^{-1}(0)\to A^{-1}(0)\quad\text{as } t-s\downarrow 0 \text{ in norm.}$$

Thus we have arrived at the following assertion:

THEOREM 2.3. *Suppose that* (1.2), (1.3), *and* (2.5) *are satisfied. Then we have*

(2.16) (i) $$\frac{\partial}{\partial t} U(t,s) = A(t)U(t,s), \qquad t > s.$$

(ii) *$U(t,s)A^{-1}(0)$ converges to $A^{-1}(0)$ in norm as $t-s \downarrow 0$.*

(iii) *For each $T>0$, there exist constants $c_9>0$ and $c_{10}>0$ such that the estimates*

(2.17) $$\|A(t)U(t,s)\| \leqq c_9(t-s)^{1-2/\beta},$$

(2.18) $$\|A(t)U(t,s)A^{-1}(s)\| \leqq c_{10}(t-s)^{1-1/\beta}$$

hold for $0 \leqq s < t \leqq T$.

Proof. The estimates (2.17) and (2.18) follow from (1.6), (2.4), (2.6), (2.7), and (2.15). Q.E.D.

Let us consider the following nonhomogeneous equation:

(2.19) $$\frac{d}{dt} x(t) = A(t)x(t) + f(t), \qquad t > s.$$

For the existence of the solution of (2.19), we have the following lemma:

LEMMA 2.4. *Suppose that $f(t)$ is Hölder continuous with exponent $\theta > 2(1/\beta - 1)$ in* (2.19). *Then the function*

$$x(t) = \int_s^t U(t,\tau) f(\tau)\, d\tau$$

gives a solution of (2.19) *with $x(s) = 0$.*

Proof. Let

$$W(t,\tau) = \int_\tau^t e^{(t-\sigma)A(\sigma)} R(\sigma,\tau)\, d\sigma,$$

and let $\delta > 0$ and $0 < \varepsilon < \delta/2$ be given. Obviously the function

(2.20) $$\frac{\partial}{\partial t} \int_s^{t-\varepsilon} W(t,\tau) f(\tau)\, d\tau = \int_s^{t-\varepsilon} \frac{\partial}{\partial t} W(t,\tau) f(\tau)\, d\tau + W(t, t-\varepsilon) f(t-\varepsilon)$$

converges to

(2.21) $$\int_s^t \frac{\partial}{\partial t} W(t,\tau) f(\tau)\, d\tau$$

uniformly in $t \in [s+\delta, T]$ as $\varepsilon \downarrow 0$. In the same way as (2.11), we find that the function

(2.22)
$$\begin{aligned} &\frac{\partial}{\partial t} \int_s^{t-\varepsilon} e^{(t-\tau)A(\tau)} f(\tau)\, d\tau \\ &= e^{\varepsilon A(t-\varepsilon)} f(t-\varepsilon) - \int_s^{t-\varepsilon} S(t,\tau) f(\tau)\, d\tau \\ &\quad + \int_s^{t-\varepsilon} A(t)\, e^{(t-\tau)A(t)} \{f(\tau) - f(t)\}\, d\tau \\ &\quad + \{e^{(t-s)A(t)} - e^{\varepsilon A(t)}\} f(t) \end{aligned}$$

converges to

$$(2.23)\qquad -\int_s^t S(t,\tau)f(\tau)\,d\tau+\int_s^t A(t)\,e^{(t-\tau)A(t)}\{f(\tau)-f(t)\}\,d\tau+e^{(t-s)A(t)}f(t)$$

uniformly in $t\in[s+\delta, T]$ as $\varepsilon\downarrow 0$. It follows from (2.20) and (2.22) that

$$(2.24)\qquad \begin{aligned} A(t)&\int_s^{t-\varepsilon} U(t,\tau)f(\tau)\,d\tau \\ &=-\int_s^{t-\varepsilon} S(t,\tau)f(\tau)\,d\tau+\int_s^{t-\varepsilon} A(t)\,e^{(t-\tau)A(t)}\{f(\tau)-f(t)\}\,d\tau \\ &\quad+\{e^{(t-s)A(t)}-e^{\varepsilon A(t)}\}f(t)+\int_s^{t-\varepsilon}\frac{\partial}{\partial t}W(t,\tau)f(\tau)\,d\tau. \end{aligned}$$

Therefore it follows from (2.20) to (2.24) that

$$A(t)\Big[\int_s^{t-\varepsilon} U(t,\tau)f(\tau)\,d\tau+e^{\varepsilon A(t)}A^{-1}(t)f(t)\Big]$$

converges to

$$(2.25)\qquad \begin{aligned} &-\int_s^t S(t,\tau)f(\tau)\,d\tau+\int_s^t A(t)\,e^{(t-\tau)A(t)}\{f(\tau)-f(t)\}\,d\tau \\ &+e^{(t-s)A(t)}f(t)+\int_s^t\frac{\partial}{\partial t}W(t,\tau)f(\tau)\,d\tau \\ &\qquad\qquad=\frac{\partial}{\partial t}\int_s^t U(t,\tau)f(\tau)\,d\tau. \end{aligned}$$

Since $A(t)$ is closed, it follows that the equation

$$(2.26)\qquad A(t)\Big[\int_s^t U(t,\tau)f(\tau)\,d\tau+A^{-1}(t)f(t)\Big]=\frac{\partial}{\partial t}\int_s^t U(t,\tau)f(\tau)\,d\tau$$

holds for $t>s$. Therefore $x(t)$ satisfies (2.19) and $x(s)=0$. Q.E.D.

By using Lemma 2.4, we obtain the following assertion:

THEOREM 2.5. *If $f(t)$ is Hölder continuous with exponent $\theta>2(1/\beta-1)$, the solution of* (2.19) *with $x(s)=x_0$ is unique.*

Proof. Let $x(t)$ be a solution of (2.19) with $x(s)=x_0$. In the way similar to [6], define $z_\varepsilon(\tau)$ as

$$z_\varepsilon(\tau)=U(\tau,s+\varepsilon)x(s+\varepsilon)+\int_{s+\varepsilon}^{\tau} U(\tau,\sigma)f(\sigma)\,d\sigma-x(\tau),\qquad \tau>s+\varepsilon,$$

$$z_\varepsilon(s+\varepsilon)=0.$$

Since $x(s+\varepsilon)\in D(A)$, $z_\varepsilon(\tau)$ is continuous for $\tau\geqq s+\varepsilon$. From Theorem 2.3 and Lemma 2.4 it follows that

$$(2.27)\qquad \begin{aligned} \frac{d}{d\tau}z_\varepsilon(\tau)&=A(\tau)z_\varepsilon(\tau) \\ &=A(t)z_\varepsilon(\tau)+[A(\tau)-A(t)]A^{-1}(\tau)A(\tau)z_\varepsilon(\tau), \end{aligned}$$

where $t>s+\varepsilon$ and $\tau>s+\varepsilon$. It is easy to see that $\|A(\tau)z_\varepsilon(\tau)\|$ is integrable on $(s+\varepsilon, T)$, where $T>s+\varepsilon$ is arbitrary. Therefore from (2.27) we obtain the integral equation

$$z_\varepsilon(t)=\int_{s+\varepsilon}^{t} e^{(t-\sigma)A(t)}[A(\sigma)-A(t)]z_\varepsilon(\sigma)\,d\sigma, \qquad t>s+\varepsilon. \tag{2.28}$$

It follows from (1.3), (1.6), and (2.28) that the inequality

$$\|A(t)z_\varepsilon(t)\|\leqq\int_{s+\varepsilon}^{t} c_{11}(t-\sigma)^{\rho+1-2/\beta}\|A(\sigma)z_\varepsilon(\sigma)\|d\sigma$$

holds for $t>s+\varepsilon$. The above inequality implies that

$$A(t)z_\varepsilon(t)=0, \qquad t>s+\varepsilon.$$

Therefore we obtain

$$U(t, s+\varepsilon)x(s+\varepsilon)+\int_{s+\varepsilon}^{t} U(t,\sigma)f(\sigma)\,d\sigma=x(t), \qquad t>s+\varepsilon.$$

Let $\varepsilon\downarrow 0$. Then we obtain

$$U(t,s)x(s)+\int_{s}^{t} U(t,\sigma)f(\sigma)\,d\sigma=x(t), \qquad t>s. \tag{2.29}$$

This completes the proof. Q.E.D.

Remark. Let $f(t)=0$. Then this theorem implies the relation

$$U(t,\tau)U(\tau,s)=U(t,s), \qquad 0\leqq s<\tau<t<+\infty. \tag{2.30}$$

If $A(t)$, $t\geqq 0$, is strongly continuously differentiable on $D(A)$, (1.3) is satisfied with $\rho=1$. Then we have the following assertion which is stronger than Theorem 2.5 and has not been obtained in [5], [13]:

THEOREM 2.6. *Suppose that $A(t)$, $t\geqq 0$, is strongly continuously differentiable on $D(A)$ and that the assumptions of Theorem 2.3 are satisfied. Then $U(t,s)A^{-1}(0)$ is strongly continuously differentiable in s $(<t)$ and the equation*

$$\frac{\partial}{\partial s}U(t,s)A^{-1}(0)=-U(t,s)A(s)A^{-1}(0) \tag{2.31}$$

holds for $s<t$.

For the proof of the above theorem, we define fractional powers of the operator $A(t)$ as follows [1]:

$$A^{-\gamma}(t)=\frac{e^{-i\pi\gamma}}{\Gamma(\gamma)}\int_0^\infty s^{\gamma-1}e^{sA(t)}\,ds, \qquad t\geqq 0, \tag{2.32}$$

where $\gamma>1/\beta-1$. If γ is an integer, it is easy to see that $A^{-\gamma}(t)$ defined in (2.32) coincides with $A^{-\gamma}(t)$ in the usual sense. It can be proved by the standard method that the inverse $A^{\gamma}(t)$ of $A^{-\gamma}(t)$ exists for $\gamma>1/\beta-1$ [1], [6]. The range of $A^{-\gamma}(t)$ will be denoted by $D(A^{\gamma}(t))$. The following relations (i) to (iv) are proved by a well-known method:

(i) If $\alpha, \gamma > 1/\beta - 1$, then

$$A^{-\alpha}(t)A^{-\gamma}(t) = A^{-(\alpha+\gamma)}(t).$$

(ii) If $\alpha, \gamma, \alpha-\gamma > 1/\beta - 1$, then

$$A^{\gamma}(t)A^{-\alpha}(t) = A^{-(\alpha-\gamma)}(t),$$
$$A^{-\alpha}(t)A^{\gamma}(t) = A^{-(\alpha-\gamma)}(t) \quad \text{on } D(A^{\gamma}(t)).$$

(iii) If $\alpha, \gamma, \gamma-\alpha > 1/\beta - 1$, then

$$A^{\gamma}(t)A^{-\alpha}(t) = A^{\gamma-\alpha}(t) \quad \text{on } D(A^{\gamma-\alpha}(t)),$$
$$A^{-\alpha}(t)A^{\gamma}(t) = A^{\gamma-\alpha}(t) \quad \text{on } D(A^{\gamma}(t)).$$

(iv) If $\alpha, \gamma > 1/\beta - 1$, then

$$A^{\alpha}(t)A^{\gamma}(t) = A^{\gamma}(t)A^{\alpha}(t) = A^{\alpha+\gamma}(t) \quad \text{on } D(A^{\alpha+\gamma}(t)).$$

Let $\Gamma_\mu = \{-\mu t;\, t>0\}$, where $\operatorname{Re}\mu < 0$. We note the equations

$$\int_{\Gamma_\mu} e^{-z}z^{\gamma-1}\,dz = \Gamma(\gamma), \quad \text{if } \operatorname{Im}\mu \leqq 0,$$

$$\int_{\Gamma_\mu} e^{-z}z^{\gamma-1}\,dz = e^{i2\pi\gamma}\Gamma(\gamma), \quad \text{if } \operatorname{Im}\mu > 0,$$

where $\gamma > 0$ and $z^{\gamma-1}$ is considered as single-valued in the plane with a cut along the positive real semiaxis, i.e., $z^{\gamma-1} = r^{\gamma-1}e^{i\theta(\gamma-1)}$ for $z = re^{i\theta}$, $0 \leqq \theta < 2\pi$. Let $\Gamma_- = \{\mu \in \Gamma_q;\, \operatorname{Im}\mu \leqq 0\}$ and let $\Gamma_+ = \{\mu \in \Gamma_q;\, \operatorname{Im}\mu > 0\}$ respectively. Then we obtain the following equations [5]:

$$\begin{aligned}
(2.33)\qquad & -\frac{1}{2\pi i}\int_{\Gamma_q} \mu^{-\gamma}(A(t)-\mu I)^{-1}\,d\mu \\
&= -\frac{1}{2\pi i}\int_{\Gamma_-} \mu^{-\gamma}\frac{1}{\Gamma(\gamma)}\int_{\Gamma_\mu} e^{-z}z^{\gamma-1}\,dz\,(A(t)-\mu I)^{-1}\,d\mu \\
&\quad -\frac{1}{2\pi i}\int_{\Gamma_+} \mu^{-\gamma}\frac{e^{-2\pi\gamma i}}{\Gamma(\gamma)}\int_{\Gamma_\mu} e^{-z}z^{\gamma-1}dz(A(t)-\mu I)^{-1}\,d\mu \\
&= -\frac{e^{-i\pi\gamma}}{2\pi i\Gamma(\gamma)}\int_{\Gamma_-}\int_0^\infty e^{\mu s}s^{\gamma-1}\,ds\,(A(t)-\mu I)^{-1}\,d\mu \\
&\quad -\frac{e^{-i\pi\gamma}}{2\pi i\Gamma(\gamma)}\int_{\Gamma_+}\int_0^\infty e^{\mu s}s^{\gamma-1}\,ds\,(A(t)-\mu I)^{-1}\,d\mu \\
&= \frac{e^{-i\pi\gamma}}{\Gamma(\gamma)}\int_0^\infty s^{\gamma-1}e^{sA(t)}\,ds,
\end{aligned}$$

where $\gamma > 1/\beta - 1$ and $\mu^{-\gamma} = r^{-\gamma}e^{-i\theta\gamma}$ for $\mu = re^{i\theta}$, $0 \leqq \theta < 2\pi$. In the same way as (2.33), we obtain

$$(2.34)\qquad A^{-\gamma}(t)\,e^{\tau A(t)} = -\frac{1}{2\pi i}\int_{\Gamma_q} \mu^{-\gamma}e^{\mu\tau}(A(t)-\mu I)^{-1}\,d\mu, \qquad \tau > 0.$$

We note that $e^{\tau A(t)}x \in D(A^n(t))$, $n=0, 1, \cdots$, for each $x \in E$. It follows easily from (2.34) that

$$A^{\gamma}(t)\, e^{\tau A(t)} = -\frac{1}{2\pi i}\int_{\Gamma_q} \mu^{\gamma} e^{\mu\tau}(A(t)-\mu I)^{-1}\, d\mu, \tag{2.35}$$

where $\gamma > 1/\beta - 1$, and $\tau > 0$. It is clear that $A^{\gamma}(t)\, e^{\tau A(t)}$ is a bounded operator and continuous in norm for $t \geqq 0$ and $\tau > 0$. It follows from (2.35) that the estimate

$$\|A^{\gamma}(t)\, e^{\tau A(t)}\| \leqq c(\gamma)\, e^{-\lambda\tau}\tau^{1-(\gamma+1)/\beta} \tag{2.36}$$

holds for $\gamma > 1/\beta - 1$, $\tau > 0$, and $t \geqq 0$. Thus we have arrived at the following lemma:

LEMMA 2.7. *$A^{\gamma}(t)e^{\tau A(t)}$ is continuous in norm for $t \geqq 0$ and $\tau > 0$. If $2\beta - 1 > \gamma > 1/\beta - 1$, $\|A^{\gamma}(t)e^{\tau A(t)}\|$ is integrable for τ on $(0, \infty)$.*

LEMMA 2.8. *If $\gamma > 2 - 2\beta$, then $A^{-\gamma}(t)$ is strongly continuously differentiable in $t \geqq 0$.*

Proof. We use the representation (2.33) of $A^{-\gamma}(t)$. From the relation $\gamma > 2 - 2\beta$, the function

$$\left\|\frac{\partial}{\partial t}\mu^{-\gamma}(A(t)-\mu I)^{-1}x\right\| = \|-\mu^{-\gamma}(A(t)-\mu I)^{-1}A'(t)(A(t)-\mu I)^{-1}x\|$$

is integrable on Γ_q. This completes the proof. Q.E.D.

LEMMA 2.9. *If $\gamma > 2 - 2\beta$, then $A^{-\gamma}(t)\, e^{(t-s)A(s)}$ converges to $A^{-\gamma}(s)$ in norm as $t \downarrow s$.*

Proof. Consider the equation

$$A^{-\gamma}(t)\, e^{(t-s)A(s)} - A^{-\gamma}(s) = -\frac{1}{2\pi i}\int_{\Gamma_q} \mu^{-\gamma}(e^{\mu(t-s)}-1)(A(s)-\mu I)^{-1}\, d\mu$$
$$+ (A^{-\gamma}(t) - A^{-\gamma}(s))\, e^{(t-s)A(s)}.$$

It is clear that the first term converges to 0 as $t \downarrow s$. From (1.6) and Lemma 2.8 it follows that the second term also converges to 0 as $t \downarrow s$. Q.E.D.

Now let us prove Theorem 2.6, using the above lemmas.

Proof of Theorem 2.6. Following [10], [12], let us define $Q_1(t, s)$ as

$$Q_1(t, s) = \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial s}\right) e^{(t-s)A(s)}$$
$$= -\frac{1}{2\pi i}\int_{\Gamma_q} e^{\mu(t-s)} \frac{\partial}{\partial s}(A(s)-\mu I)^{-1}\, d\mu.$$

$Q_1(t, s)$ is strongly continuous for $t > s$ and the estimate

$$\|Q_1(t, s)\| \leqq c_{12}(t-s)^{2(1-1/\beta)}$$

holds for $0 \leqq s < t \leqq T$. Let $Q(t, s)$ be the strongly continuous operator which satisfies the integral equation

$$Q(t, s) = Q_1(t, s) + \int_s^t Q(t, \sigma)Q_1(\sigma, s)\, d\sigma.$$

Define the strongly continuous operator $V(t, s)$ as follows:

$$V(t, s) = e^{(t-s)A(s)} + \int_s^t Q(t, \tau)\, e^{(\tau-s)A(s)}\, d\tau.$$

It is easy to see that $V(t, s)$ is strongly continuously differentiable in s on $D(A)$ and satisfies the equation

$$\frac{\partial}{\partial s} V(t,s)A^{-1}(0)x + V(t,s)A(s)A^{-1}(0)x = 0 \tag{2.37}$$

for $t>s$ and for each $x\in E$. It follows from (2.16) and (2.37) that the function $V(t,r)U(r,s)x$ is independent of r, $s<r<t$, for each $x\in E$. Since $V(t,r)A^{-1}(r)$ converges to $A^{-1}(t)$ in norm and $A(r)U(r,s)$ converges to $A(t)U(t,s)$ in norm respectively as $r\uparrow t$, it follows that

$$V(t,r)U(r,s)x \to A^{-1}(t)A(t)U(t,s)x = U(t,s)x \quad \text{as } r\uparrow t.$$

Choose $\gamma>0$ satisfying $2-2\beta<\gamma<2\beta-1$, and consider the equation

$$A^{-\gamma}(r)U(r,s) = A^{-\gamma}(r)\, e^{(r-s)A(s)} + \int_s^r A^{-\gamma}(r)\, e^{(r-\tau)A(\tau)}R(\tau,s)\, d\tau.$$

Since $A^{-\gamma}(r)\, e^{(r-\tau)A(\tau)}$ is uniformly bounded for $r>\tau$, the second term converges to 0 in norm as $r\downarrow s$. Therefore it follows from Lemma 2.9 that $A^{-\gamma}(r)U(r,s)$ converges to $A^{-\gamma}(s)$ in norm as $r\downarrow s$. Consider the equation

$$V(t,r)U(r,s)x = P(t,r)A^{-\gamma}(r)U(r,s)x, \qquad s<r<t,$$

where $P(t,r)$, $t>r$, is the strongly continuous operator given by

$$P(t,r) = A^{\gamma}(r)\, e^{(t-r)A(r)} + \int_r^t Q(t,\tau)A^{\gamma}(r)\, e^{(\tau-r)A(r)}\, d\tau. \tag{2.38}$$

Here we used the fact that $e^{hA(r)}$ and $A^{\gamma}(r)$ commute on $D(A^{\gamma}(r))$.

Let $r\downarrow s$. Then $V(t,r)U(r,s)x$ converges to

$$P(t,s)A^{-\gamma}(s)x = V(t,s)x.$$

Therefore it follows that $V(t,s)=U(t,s)$ for $t>s$, and this shows that the equation (2.31) holds. Q.E.D.

3. Existence and uniqueness of solutions of (1.1). In this section we consider the problem of the existence and the uniqueness of the solution of (1.1) under the condition (1.8) for $f(t,x)$. Throughout this section, it is assumed that the conditions (1.2), (1.3), and (2.5) are satisfied. Let us consider the following integral equation corresponding to (1.1):

$$x(t) = U(t,0)x_0 + \int_0^t U(t,\tau)f(\tau, x(\tau))\, d\tau. \tag{3.1}$$

Equation (3.1) can be uniquely solved by the successive approximations:

$$\begin{aligned} &x_0(t) = U(t,0)x_0,\\ &x_{n+1}(t) = U(t,0) + \int_0^t U(t,\tau)f(\tau, x_n(\tau))\, d\tau, \qquad n=0,1,\cdots, \end{aligned} \tag{3.2}$$

$$x(t) = \lim_{n\to\infty} x_n(t). \tag{3.3}$$

If $x_0 \in D(A)$, the solution $x(t)$ of (3.1) is in $C([0, \infty); E)$. Otherwise $x(t)$ is in $C((0, \infty); E)$ and satisfies the following estimate in the neighborhood of $t=0$:

$$\|x(t)\| \leqq c_{13} t^{1-1/\beta}.$$

In addition to (1.8), suppose that $f(t, x)$ satisfies the estimate

$$\|f(t, x)-f(\tau, x)\| \leqq C_{R,T}(t-\tau)^{\theta}, \qquad 0 \leqq \tau \leqq t \leqq T, \quad \|x\| \leqq R \tag{3.4}$$

for each $T>0$ and $R>0$, where $\theta > 2(1/\beta - 1)$. Then, we obtain the following basic assertion:

THEOREM 3.1. *Suppose that $f(t, x)$ satisfies* (1.8) *and* (3.4). *Then the solution $x(t)$ of* (3.1) *with $x(0)=x_0 \in D(A)$ satisfies* (1.1). *The solution of* (1.1) *which is Hölder continuous with exponent $\nu > 2(1/\beta - 1)$ is unique.*

Proof. The uniqueness follows from Theorem 2.5. Let $x(t)$ be the solution of (3.1) with $x(0)=x_0 \in D(A)$ and let $0<\tau<t \leqq T$. Consider the equation

$$\begin{aligned} x(t)-x(\tau) = [U(t, 0)x_0 - U(\tau, 0)x_0] + \int_\tau^t U(t, \sigma) f(\sigma, x(\sigma))\, d\sigma \\ + \int_0^\tau [U(t, \sigma) - U(\tau, \sigma)] f(\sigma, x(\sigma))\, d\sigma. \end{aligned} \tag{3.5}$$

It follows from (2.18) that the estimate

$$\|U(t, 0)x_0 - U(\tau, 0)x_0\| \leqq c_{14}(t-\tau)^{2-1/\beta} \tag{3.6}$$

holds for $0<\tau<t \leqq T$. It is easy to see that the estimate

$$\left\| \int_\tau^t U(t, \sigma) f(\sigma, x(\sigma))\, d\sigma \right\| \leqq c_{15}(t-\tau)^{2-1/\beta} \tag{3.7}$$

holds for $0<\tau<t \leqq T$. It follows from (2.17) that the estimate

$$\begin{aligned} &\|U(t, \sigma) - U(\tau, \sigma)\| \\ &\quad \leqq \int_\tau^t \|A(z) U(z, \sigma)\|\, dz \leqq c_{16}(t-\tau)^{2(1/\beta-1)} (\tau-\sigma)^{-4(1/\beta-1)} \end{aligned}$$

holds for $0 \leqq \sigma < \tau < t \leqq T$. Therefore the above inequality implies that the estimate

$$\begin{aligned} \left\| \int_0^\tau [U(t, \sigma) - U(\tau, \sigma)] f(\sigma, x(\sigma))\, d\sigma \right\| \\ \leqq c_{17}(t-\tau)^{2(1/\beta-1)} \tau^{5-4/\beta} \end{aligned} \tag{3.8}$$

holds for $0<\tau<t \leqq T$. It follows from (3.6) to (3.8) that

$$\|x(t)-x(\tau)\| \leqq c_{18}(t-\tau)^{2(1/\beta-1)}, \qquad 0 \leqq \tau \leqq t \leqq T.$$

Therefore the above estimate implies that the function $g(t)=f(t, x(t))$ is Hölder continuous with exponent $2(1/\beta-1)$. Again we estimate the third term of (3.5). Consider the inequality

$$\begin{aligned}
&\left\|\int_0^\tau [U(t,\sigma)-U(\tau,\sigma)]g(\sigma)\,d\sigma\right\| \\
&\leqq \int_0^\tau \|g(\sigma)\|\,d\sigma \int_\tau^t c_9(z-\sigma)^{1-2/\beta}\,dz \\
&= \int_\tau^t dz \int_0^\tau c_9\|g(\sigma)\|(z-\sigma)^{1-2/\beta}\,d\sigma \qquad (3.9)\\
&\leqq c_9 \int_\tau^t dz \int_0^\tau |\,\|g(z)\|-\|g(\sigma)\|\,|(z-\sigma)^{1-2/\beta}\,d\sigma \\
&\quad + c_9 \int_\tau^t \|g(z)\|\,dz \int_0^\tau (z-\sigma)^{1-2/\beta}\,d\sigma.
\end{aligned}$$

For any $\varepsilon>0$ there exists $c_\varepsilon>0$ such that the estimate

$$\text{the first term of (3.9)} \leqq c_\varepsilon (t-\tau)^{1-\varepsilon} \qquad (3.10)$$

holds for $0<\tau<t\leqq T$. It is easy to see that the estimate

$$\text{the second term of (3.9)} \leqq c_{19}(t-\tau)^{3-2/\beta} \qquad (3.11)$$

holds for $0<\tau<t\leqq T$. Choose $\varepsilon>0$ sufficiently small. Then it follows from (3.6), (3.7), (3.10), and (3.11) that the estimate

$$\|x(t)-x(\tau)\| \leqq c_{20}(t-\tau)^{3-2/\beta} \qquad (3.12)$$

holds for $0\leqq\tau\leqq t\leqq T$. Therefore it follows that $f(t, x(t))$ is Hölder continuous with exponent $\nu=\min(\theta, 3-2/\beta)>2(1/\beta-1)$. The assertion of this theorem follows from Lemma 2.4. Q.E.D.

Remark. It is easy to see that the estimate (3.12) holds for the solution $x(t)$ of (3.1) with $x(0)\in D(A)$ and $\theta>0$ instead of $\theta>2(1/\beta-1)$ in (3.4).

Now let us assume that $A(t)$, $t\geqq 0$, is strongly continuously differentiable on $D(A)$. Then it follows from Theorem 2.6 that the solution of (1.1) is unique. Suppose that $f(t, x)$ satisfies the following additional conditions:

(3.13) (i) $\dfrac{\partial}{\partial t} f(t,x)=f_t(t,x)$ is continuous for $(t,x)\in[0,\infty)\times E$.

(3.14) (ii) For each $t\geqq 0$, $f(t, x)$ is Fréchet differentiable in x [12], i.e.,

$$f(t, x+z)=f(t,x)+(Df(t,x)+\bar{D}f(t,x))\cdot z+o(\|z\|; x)$$

holds when $\|z\|\to 0$, where $Df(t,x)$ (resp. $\bar{D}f(t,x)$) is a linear (resp. antilinear) bounded operator in E.

(3.15) (iii) $Df(t,x)$ and $\bar{D}f(t,x)$ are strongly continuous for $(t,x)\in[0,\infty)\times E$.

(3.16) (iv) For t, x, satisfying $0\leqq t\leqq c$, $\|x\|\leqq c$, $\|f_t(t,x)\|$, $\|Df(t,x)\|_{L(E,E)}$, and $\|\bar{D}f(t,x)\|_{L(E,E)}$ are uniformly bounded.

The following assertion will be used in § 6:

THEOREM 3.2. *Suppose that $A(t)$, $t \geqq 0$, is strongly continuously differentiable on $D(A)$, and that $f(t, x)$ satisfies* (3.13) *to* (3.16). *Then for each $x_0 \in D(A)$ there exists the unique solution $x(t)$ of* (1.1) *with $x(0) = x_0$ and the equation*

$$\begin{aligned} &\frac{d}{dt} x(t) \\ (3.17) \quad &= V(t, 0)[A(0)x_0 + f(0, x_0)] + \int_0^t V(t, \sigma)\Big\{-A'(\sigma)A^{-1}(\sigma)f(\sigma, x(\sigma)) \\ &\quad + f_\sigma(\sigma, x(\sigma)) + Df(\sigma, x(\sigma)) \cdot \frac{d}{d\sigma} x(\sigma) + \bar{D}f(\sigma, x(\sigma)) \cdot \frac{d}{d\sigma} x(\sigma)\Big\}\, d\sigma \end{aligned}$$

holds for $t > 0$, where $V(t, s) = A(t)U(t, s)A^{-1}(s)$.

Proof. It follows from (2.18), (2.25), (3.13), (3.16), and Theorem 3.1 that $dx(t)/dt$ satisfies the estimate

$$\left\|\frac{d}{dt} x(t)\right\| \leqq c_{21} t^{1-1/\beta}, \qquad 0 < t \leqq T.$$

Therefore we find that the function

$$\frac{d}{dt} f(t, x(t)) = f_t(t, x(t)) + Df(t, x(t)) \cdot \frac{d}{dt} x(t) + \bar{D}f(t, x(t)) \cdot \frac{d}{dt} x(t)$$

is continuous and integrable on $(0, T)$. This shows that the equation (3.17) holds. Q.E.D.

Remark. The above theorem can be also proved by a method similar to [9], [12], without using Theorem 3.1.

Remarks on §§ 2 *and* 3. Theorems 2.3 and 2.5 are extensions of the results of H. Tanabe [10] to the weakened Cauchy problem, which are obtained by formally letting $\beta = 1$. In the uniformly well-posed Cauchy problem, similar results to Theorem 3.1 were obtained by several authors, for example, by T. Kato [3] and A. Pazy [7].

4. Asymptotic behavior I. In this section we consider the asymptotic behavior of the solution of (3.1) under the condition (1.8) for $f(t, x)$. In addition to (1.2) and (2.5), we assume that the following conditions are satisfied:

$$(4.1) \quad \begin{aligned} &\|(A(t) - A(\tau))A^{-1}(0)\| \leqq L_1(s)(t - \tau)^\rho, \qquad s \leqq \tau \leqq t < +\infty, \\ &\lim_{s \to \infty} L_1(s) = 0, \qquad L_1(0) = L_1, \end{aligned}$$

$$(4.2) \quad \sup_{0 \leqq \tau, t} \|A(t)A^{-1}(\tau)\| = L_2 < +\infty.$$

Remark. The condition (4.1) does not guarantee the convergence of $A(t)A^{-1}(0)$ as $t \to \infty$.

The following lemma is easily proved:

LEMMA 4.1. *Suppose that $f(t)$ is a real valued continuous function of $t > 0$ and is bounded on* $(0, 1]$. *If $\sup_{t>0} f(t) = +\infty$, then we can choose a sequence $\{t_n\}_{n=m}^{\infty}$ such that*

$$f(t) \leqq f(t_n) = n,$$

$$0 < t \leqq t_n, \quad 1 < t_m < t_{m+1} < \cdots < t_n < \cdots, \quad \lim_{n \to \infty} t_n = +\infty,$$

where $m > \sup_{0 < t \leqq 1} f(t)$.

Proof. Let us define $S_n = \{t > 0; f(t) \geqq n\}$. By the assumptions, the sets S_n are closed and nonvoid. Further let $t_n = \inf t$, where the infimum is taken over S_n. Then, it is easy to see that $\{t_n\}_{n=m}^{\infty}$ is the sequence stated in this lemma. Q.E.D.

Let us prove the following theorem which is important in applications:

THEOREM 4.2. *Suppose that the conditions* (1.2), (2.5), (4.1), *and* (4.2) *are satisfied*:

(i) *For each* θ, $0 < \theta < \lambda$, *there exists* $M_\theta > 0$ *such that the estimate*

$$\|U(t, s)\| \leqq M_\theta e^{-\theta(t-s)}(t-s)^{1-1/\beta}, \qquad 0 \leqq s < t < +\infty, \tag{4.3}$$

holds.

(ii) *For each* $\varepsilon > 0$ *and* $\omega < \lambda$, *there exists* $s(\varepsilon, \omega) > 0$ *such that the estimate*

$$\int_s^t e^{\omega(t-\tau)}\|U(t, \tau)\| \, d\tau \leqq M_1(\lambda - \omega)^{-(2-1/\beta)}\Gamma(2 - 1/\beta) + \varepsilon, \tag{4.4}$$

$$s(\varepsilon, \omega) \leqq s < t < +\infty$$

holds.

Proof. (i) It follows from (2.3) that

$$\|R(t, s)\| \leqq L_1(s)L_2M_2(t-s)^{\rho+1-2/\beta} e^{-\lambda(t-s)} + \int_s^t L_1(s)L_2M_2(t-\tau)^{\rho+1-2/\beta} e^{-\lambda(t-\tau)}\|R(\tau, s)\| \, d\tau.$$

Choose γ satisfying $\theta < \gamma < \lambda$. Then we have from the above inequality

$$\begin{aligned} &\int_s^T e^{\gamma(t-s)}\|R(t, s)\| \, dt \\ &\leqq (\lambda - \gamma)^{\alpha-1}\Gamma(1-\alpha)L_1(s)L_2M_2 \\ &\quad + \int_s^T e^{\gamma(\tau-s)}\|R(\tau, s)\| \, d\tau \int_\tau^T L_1(s)L_2M_2(t-\tau)^{-\alpha} e^{-(\lambda-\gamma)(t-\tau)} \, dt \end{aligned} \tag{4.5}$$

for $T > s$, where $\alpha = 2/\beta - \rho - 1$. Choose $s(\gamma, \delta)$ so that

$$(\lambda - \gamma)^{\alpha-1}\Gamma(1-\alpha)L_1(s)L_2M_2 < \delta < 1, \qquad s \geqq s(\gamma, \delta).$$

Then it follows from (4.5) that the estimate

$$\int_s^t e^{\gamma(\tau-s)}\|R(\tau, s)\| \, d\tau < \delta(1-\delta)^{-1} \tag{4.6}$$

holds for $s(\gamma, \delta) \leqq s < t < +\infty$. Therefore from (4.6) we obtain

$$\begin{aligned} &e^{\gamma(t-s)}(t-s)^{\alpha}\|R(t, s)\| \\ &\leqq L(s)(1-\delta)^{-1} \\ &\quad + \int_s^t L(s)(t-\tau)^{-\alpha} e^{-(\lambda-\gamma)(t-\tau)} \cdot e^{\gamma(\tau-s)}(\tau-s)^{\alpha}\|R(\tau, s)\| \, d\tau, \end{aligned} \tag{4.7}$$

$$s(\gamma, \delta) \leqq s < t < +\infty,$$

where $L(s) = L_1(s)L_2M_2$. Let $f(t, s) = e^{\gamma(t-s)}(t-s)^{\alpha}\|R(t, s)\|$. Then, $f(t, s)$ is continuous for $t > s$ and $\sup_{s<t\leqq s+1} f(t, s) < +\infty$ for each s. Therefore it follows from Lemma 4.1 and (4.7) that for each $s \geqq s(\gamma, \delta)$, $f(t, s)$ is bounded for t on the interval (s, ∞).

Therefore (4.7) implies

$$\sup_t f(t,s) \leqq L(s)(1-\delta)^{-2}, \qquad s(\gamma,\delta) \leqq s.$$

Since the right-hand side of the above inequality is bounded for $s \geqq s(\gamma,\delta)$, it follows that the estimate

$$\|R(t,s)\| \leqq c_{22}(t-s)^{\rho+1-2/\beta} e^{-\gamma(t-s)} \tag{4.8}$$

holds for $s(\gamma,\delta) \leqq s < t < +\infty$, where $c_{22} = (1-\delta)^{-2} \sup_{s(\gamma,\delta)\leqq s} L(s)$. The above inequality and (2.2) imply that

$$\|U(t,s)\| \leqq c_{23}(t-s)^{1-1/\beta} e^{-\theta(t-s)}, \qquad s(\gamma,\delta) \leqq s < t < +\infty. \tag{4.9}$$

Let $s_0 = s(\gamma,\delta)$ and take s_1 and s_* satisfying $s_0 < s_* < s_1$. We estimate $\|U(t,s)\|$ for $t \geqq s_1$ and $0 \leqq s \leqq s_0$. We note the inequality

$$\|U(\tau,\sigma)\| \leqq c_{24}(\tau-\sigma)^{1-1/\beta}, \qquad 0 \leqq \sigma < \tau \leqq s_1. \tag{4.10}$$

It follows from (4.9) and (4.10) that

$$\begin{aligned} \|U(t,s)\| &\leqq \|U(t,s_*)\| \, \|U(s_*,s)\| \\ &\leqq c_{23}c_{24} e^{-\theta(t-s_*)}(t-s_*)^{1-1/\beta}(s_*-s)^{1-1/\beta} \\ &\leqq c_{23}c_{24} e^{\theta s_*}\left(\frac{s_1-s_0}{(s_1-s_*)(s_*-s_0)}\right)^{1/\beta-1} e^{-\theta(t-s)}(t-s)^{1-1/\beta}, \\ &\qquad\qquad 0 \leqq s \leqq s_0, \qquad s_1 \leqq t. \end{aligned} \tag{4.11}$$

Therefore (4.9) to (4.11) implies (4.3).

(ii) Set

$$g(t,s) = \int_s^t e^{\omega(t-\tau)} \|R(t,\tau)\| \, d\tau, \qquad t > s.$$

$g(t,s)$ is continuous for $t > s$ and tends to 0 as $t \downarrow s$. In the same way as (4.5), we obtain

$$\begin{aligned} g(t,s) &\leqq L(s)(\lambda-\omega)^{\alpha-1}\Gamma(1-\alpha) \\ &\quad + \int_s^t L(s) e^{-(\lambda-\omega)(t-\sigma)}(t-\sigma)^{-\alpha} g(\sigma,s) \, d\sigma. \end{aligned} \tag{4.12}$$

Choose s_0 so large that $L(s)(\lambda-\omega)^{\alpha-1}\Gamma(1-\alpha) < 1$ for $s \geqq s_0$. Then in the same way as in (i), (4.12) implies the estimate

$$g(t,s) \leqq \frac{L(s)(\lambda-\omega)^{\alpha-1}\Gamma(1-\alpha)}{1-L(s)(\lambda-\omega)^{\alpha-1}\Gamma(1-\alpha)}, \qquad s_0 \leqq s \leqq t < +\infty. \tag{4.13}$$

Consider the inequality

$$
\begin{aligned}
&\int_s^t e^{\omega(t-\tau)}\|U(t,\tau)\|\,d\tau \\
&\quad\leqq \int_s^t e^{\omega(t-\tau)}\|e^{(t-\tau)A(\tau)}\|\,d\tau \\
&\qquad + \int_s^t d\tau \int_\tau^t e^{\omega(t-\sigma)}\|e^{(t-\sigma)A(\sigma)}\|\, e^{\omega(\sigma-\tau)}\|R(\sigma,\tau)\|\,d\sigma \qquad (4.14)\\
&\quad\leqq M_1(\lambda-\omega)^{-(2-1/\beta)}\Gamma(2-1/\beta) \\
&\qquad + \int_s^t M_1\, e^{-(\lambda-\omega)(t-\sigma)}(t-\sigma)^{1-1/\beta} g(\sigma,s)\,d\sigma.
\end{aligned}
$$

Expression (4.4) immediately follows from (4.13) and (4.14). Q.E.D.

The following theorem is also important in applications and will be used in § 6:

THEOREM 4.3. *Suppose that the conditions* (1.2), (2.5), (4.1), *and* (4.2) *are satisfied.*

(i) *For each* θ, $0<\theta<\lambda$, *there exists* $\tilde{M}_\theta>0$ *such that the estimate*

$$\|V(t,s)\|\leqq \tilde{M}_\theta\, e^{-\theta(t-s)}(t-s)^{1-1/\beta}, \qquad 0\leqq s<t<+\infty \tag{4.15}$$

holds.

(ii) *For each* $\omega<\lambda$, *there exists* $s(\omega)>0$ *such that the estimates*

$$
\begin{aligned}
&\int_s^t e^{\omega(t-\tau)}\|V(t,\tau)\|\,d\tau \\
&\quad\leqq \frac{M_1L_2(s)(\lambda-\omega)^{-(2-1/\beta)}\Gamma(2-1/\beta)}{1-L_1(s)L_2M_2(\lambda-\omega)^{-(\rho+2-2/\beta)}\Gamma(\rho+2-2/\beta)}, \qquad s(\omega)\leqq s<t<+\infty,
\end{aligned} \tag{4.16}
$$

$$
\begin{aligned}
&\int_s^T e^{\omega(t-s)}\|V(t,s)\|\,dt \\
&\quad\leqq \frac{M_1L_2(s)(\lambda-\omega)^{-(2-1/\beta)}\Gamma(2-1/\beta)}{1-L_1(s)L_2M_2(\lambda-\omega)^{-(\rho+2-2/\beta)}\Gamma(\rho+2-2/\beta)}, \qquad s(\omega)\leqq s<T<+\infty
\end{aligned} \tag{4.17}
$$

hold, where $V(t,s)$ *is defined in Theorem* 3.2, *and* $L_2(s)$ *is given by*

$$L_2(s)=\sup_{s\leqq\tau,t}\|A(t)A^{-1}(\tau)\|. \tag{4.18}$$

Proof. (i) It is easy to see that $V(t,s)$ satisfies the integral equation [6], [1]

$$
\begin{aligned}
V(t,s)&=A(t)\,e^{(t-s)A(t)}A^{-1}(s) \\
&\quad+\int_s^t A(t)\,e^{(t-\tau)A(t)}[A(\tau)-A(t)]A^{-1}(\tau)V(\tau,s)\,d\tau.
\end{aligned} \tag{4.19}
$$

Hence it follows from (4.19) that

$$
\begin{aligned}
\|V(t,s)\|&\leqq M_1L_2(s)\,e^{-\lambda(t-s)}(t-s)^{1-1/\beta} \\
&\quad+\int_s^t L(s)\,e^{-\lambda(t-\tau)}(t-\tau)^{\rho+1-2/\beta}\|V(\tau,s)\|\,d\tau.
\end{aligned} \tag{4.20}
$$

In the same way as (4.6), for each δ, $0<\delta<1$, there exists $s(\theta,\delta)>0$ such that

$$\int_s^t e^{\theta(\tau-s)}\|V(\tau,s)\|\,d\tau \leqq (1-\delta)^{-1}M_1L_2(s)(\lambda-\theta)^{-(2-1/\beta)}\Gamma(2-1/\beta), \qquad s(\theta,\delta)\leqq s\leqq t<+\infty. \tag{4.21}$$

Thus, from (4.20) we obtain the inequality

$$\begin{aligned} &e^{\theta(t-s)}(t-s)^{-(\rho+1-2/\beta)}\|V(t,s)\| \\ &\leqq M_1L_2(s)\left(\frac{-(\rho-1/\beta)}{(\lambda-\theta)e}\right)^{-(\rho-1/\beta)} \\ &\quad +L(s)(1-\delta)^{-1}M_1L_2(s)(\lambda-\theta)^{-(2-1/\beta)}\Gamma(2-1/\beta) \\ &\quad +\int_s^t L(s)\,e^{-(\lambda-\theta)(t-\tau)}(t-\tau)^{\rho+1-2/\beta}\,e^{\theta(\tau-s)}(\tau-s)^{-(\rho+1-2/\beta)}\|V(\tau,s)\|\,d\tau, \\ &\qquad s(\theta,\delta)\leqq s<t<+\infty. \end{aligned} \tag{4.22}$$

Let $h(t,s)=e^{\theta(t-s)}(t-s)^{-(\rho+1-2/\beta)}\|V(t,s)\|$. Then $h(t,s)$ is continuous for $t>s$ and tends to 0 as $t\downarrow s$. Therefore, in the same way as Theorem 4.2, the estimate (4.22) implies

$$\begin{aligned} \sup_{s(\theta,\delta)\leqq s<t} h(t,s)\leqq (1-\delta)^{-1}M_1\Bigg\{&\left(\frac{-(\rho-1/\beta)}{(\lambda-\theta)e}\right)^{-(\rho-1/\beta)}\cdot\sup_{s(\theta,\delta)\leqq s}L_2(s) \\ &+(1-\delta)^{-1}(\lambda-\theta)^{-(2-1/\beta)}\Gamma(2-1/\beta)\cdot\sup_{s(\theta,\delta)\leqq s}L(s)L_2(s)\Bigg\}. \end{aligned}$$

Therefore the above inequality implies that the estimate

$$\|V(t,s)\|\leqq c_{25}\,e^{-\theta(t-s)}(t-s)^{1-1/\beta} \tag{4.23}$$

holds for $s(\theta,\delta)\leqq s$ and $s+1\leqq t$, where $c_{25}=(1-\delta)^{-1}M_1\{\cdot\}$.

We consider the inequality

$$\begin{aligned} &(t-s)^{-(1-1/\beta)}\|V(t,s)\| \\ &\leqq M_1L_2(s)\,e^{-\lambda(t-s)}+\int_s^t L(s)\,e^{-\lambda(t-\tau)}(t-\tau)^{\rho-1/\beta}\|V(\tau,s)\|\,d\tau \\ &\quad +\int_s^t L(s)\,e^{-\lambda(t-\tau)}(t-\tau)^{\rho+1-2/\beta}(\tau-s)^{-(1-1/\beta)}\|V(\tau,s)\|\,d\tau \\ &\leqq M_1L_2(s)+L(s)\{B(\rho+1-1/\beta,2-1/\beta)+(\rho+2-2/\beta)^{-1}\} \\ &\quad\cdot\sup_\tau(\tau-s)^{-(1-1/\beta)}\|V(\tau,s)\|, \\ &\qquad s<t<s+1, \end{aligned} \tag{4.24}$$

where the supremum in τ is taken over the interval $(s,s+1)$. Therefore, from (4.24) there exists $s_0>0$ such that

$$\begin{aligned} \|V(t,s)\|&\leqq c_{26}(t-s)^{1-1/\beta} \\ &\leqq c_{26}\,e^{\theta}\cdot e^{-\theta(t-s)}(t-s)^{1-1/\beta}, \qquad s_0\leqq s<t\leqq s+1<+\infty. \end{aligned} \tag{4.25}$$

Inequalities (4.23) and (4.25) imply that

$$\|V(t, s)\| \leqq \max\{c_{25}, c_{26}\, e^{\theta}\}\, e^{-\theta(t-s)}(t-s)^{1-1/\beta},$$

$$\max\{s(\theta, \delta), s_0\} \leqq s < t < +\infty.$$

We note the equation

$$V(t, \tau)V(\tau, s) = V(t, s), \qquad s < \tau < t.$$

Then, the rest of the proof can be carried out in the same way as Theorem 4.2, (i).

(ii) The proofs of (4.16) and (4.17) are the same as (4.13) and (4.6) respectively. Hence we omit them. Q.E.D.

Now let us consider the asymptotic behavior of the solution of (3.1).

THEOREM 4.4. *Suppose that*

$$KM_1\lambda^{-(2-1/\beta)}\Gamma(2-1/\beta) < 1. \tag{4.26}$$

Let $x(t)$ *be the solution of* (3.1). *Then we have*

$$\overline{\lim_{t\to\infty}}\, \|x(t)\| \leqq \left(\frac{M_1\lambda^{-(2-1/\beta)}\Gamma(2-1/\beta)}{1-KM_1\lambda^{-(2-1/\beta)}\Gamma(2-1/\beta)}\right) \overline{\lim_{t\to\infty}}\, \|f(t, 0)\|. \tag{4.27}$$

Proof. If $\overline{\lim}_{t\to\infty}\|f(t, 0)\| = +\infty$, then (4.27) is clear. Therefore we assume henceforth that $\overline{\lim}_{t\to\infty}\|f(t, 0)\| < +\infty$. We note the equation

$$x(t) = U(t, s)x(s) + \int_s^t U(t, \tau)f(\tau, x(\tau))\, d\tau, \qquad 0 \leqq s < t < +\infty.$$

Take any $\varepsilon > 0$ satisfying

$$K(M_1\lambda^{-(2-1/\beta)}\Gamma(2-1/\beta) + \varepsilon) < 1.$$

By (4.4) we can choose $s(\varepsilon) > 0$ so that the inequality

$$\int_s^t K\|U(t, \tau)\|\, d\tau \leqq K(M_1\lambda^{-(2-1/\beta)}\Gamma(2-1/\beta) + \varepsilon) < 1 \tag{4.28}$$

holds for $s(\varepsilon) \leqq s \leqq t < +\infty$. Consider the inequality

$$\begin{aligned} \|x(t)\| \leqq \|U(t, s)x(s)\| &+ \int_s^t \|U(t, \tau)\|\, \|f(\tau, 0)\|\, d\tau \\ &+ \int_s^t K\|U(t, \tau)\|\, \|x(\tau)\|\, d\tau, \quad s(\varepsilon) \leqq s < t. \end{aligned} \tag{4.29}$$

Since the first and the second terms of (4.29) are bounded for $t(>s)$, it follows from (4.28) and Lemma 4.1 that $\|x(t)\|$ is bounded for $t(\geqq s)$. If necessary, choose $s(\varepsilon)$ so large that the inequalities

$$\sup_{\tau \geqq s(\varepsilon)} \|x(\tau)\| \leqq \overline{\lim_{t\to\infty}}\, \|x(t)\| + \varepsilon, \qquad \sup_{\tau \geqq s(\varepsilon)} \|f(\tau, 0)\| \leqq \overline{\lim_{t\to\infty}}\, \|f(t, 0)\| + \varepsilon$$

hold. Then we obtain

$$\begin{aligned} \|x(t)\| \leqq \|U(t, s)x(s)\| &+ \left(\overline{\lim_{t\to\infty}}\, \|f(t, 0)\| + \varepsilon\right)(M_1\lambda^{-(2-1/\beta)}\Gamma(2-1/\beta) + \varepsilon) \\ &+ \left(\overline{\lim_{t\to\infty}}\, \|x(t)\| + \varepsilon\right)K(M_1\lambda^{-(2-1/\beta)}\Gamma(2-1/\beta) + \varepsilon), \qquad s(\varepsilon) \leqq s < t < +\infty. \end{aligned} \tag{4.30}$$

Therefore (4.30) and (4.3) imply (4.27). Q.E.D.

If $f(t, 0)$ tends to 0 when $t \to \infty$, the following theorem holds:

THEOREM 4.5. *Suppose that* (4.26) *is satisfied, and let* $x(t)$ *be the solution of* (3.1).

(i) *If there exist* $c>0$ *and* $\delta>0$ *such that the estimate*

$$\|f(t, 0)\| \leqq c\, e^{-\delta t}, \qquad t \geqq 0, \tag{4.31}$$

holds, then $x(t)$ *satisfies the estimate*

$$\|x(t)\| \leqq \tilde{c}\, e^{-\tilde{\delta} t} t^{1-1/\beta}, \qquad t>0, \tag{4.32}$$

where $\tilde{c}>0$ *and* $\tilde{\delta}>0$ *are some constants.*

(ii) *If there exist* $c>0$, *an integer* $n \geqq 0$, *and* β, $0 \leqq \beta < 1$, *such that the estimate*

$$\|f(t, 0)\| \leqq c\, t^{-(n+\beta)} \tag{4.33}$$

holds when $t \to \infty$, *then* $x(t)$ *satisfies the estimate*

$$\|x(t)\| \leqq \tilde{c}\, t^{-(n+\beta)} \tag{4.34}$$

when $t \to \infty$, *where* $\tilde{c}>0$ *is some constant.*

(iii) *If there exists* $c>0$ *such that the estimate*

$$\|f(t, 0)\| \leqq c\{\ln t\}^{-1} \tag{4.35}$$

holds when $t \to \infty$, *then* $x(t)$ *satisfies the estimate*

$$\|x(t)\| \leqq \tilde{c}\{\ln t\}^{-1} \tag{4.36}$$

when $t \to \infty$, *where* $\tilde{c}>0$ *is some constant.*

Proof. (i) Choose $\omega>0$ such that

$$\omega < \min(\delta, \lambda), \qquad KM_1(\lambda-\omega)^{-(2-1/\beta)}\Gamma(2-1/\beta)<1.$$

Consider the inequality

$$\begin{aligned} e^{\omega t}\|x(t)\| \leqq e^{\omega t}\|U(t, s)x(s)\| &+ \int_s^t e^{\omega(t-\tau)}\|U(t, \tau)\|\, e^{\omega\tau}\|f(\tau, 0)\|\, d\tau \\ &+ \int_s^t e^{\omega(t-\tau)} K\|U(t, \tau)\|\, e^{\omega\tau}\|x(\tau)\|\, d\tau. \end{aligned} \tag{4.37}$$

(4.3) and (4.31) imply that the first and the second terms of (4.37) are bounded for $t(>s)$ and tend to 0 exponentially as $t \to \infty$ for each $s>0$. Take $s(\omega)>0$ so that the estimate

$$\int_s^t e^{\omega(t-\tau)} K\|U(t, \tau)\|\, d\tau < 1$$

holds for $s(\omega) \leqq s < t < +\infty$. Then, in the same way as in Theorem 4.4 we find that $e^{\omega t}\|x(t)\|$ is bounded for $t \geqq 1$, and this implies (4.32).

(ii) Let $n \geqq 1$. We consider the inequality

$$\begin{aligned} t^{n+\beta} &\int_s^t \|U(t, \tau)\|\, \|f(\tau, 0)\|\, d\tau \\ &\leqq \int_s^t \|U(t, \tau)\| \sum_{j=0}^{n} {}_nC_j (t-\tau)^{n-j+\beta}\tau^j \|f(\tau, 0)\|\, d\tau \\ &\quad + \int_s^t \|U(t, \tau)\| \sum_{j=0}^{n} {}_nC_j (t-\tau)^{n-j}\tau^{j+\beta} \|f(\tau, 0)\|\, d\tau. \end{aligned}$$

It follows from (4.3) and (4.33) that the right-hand side of the above inequality is bounded for $t>s\geqq 0$. Next consider the inequality

$$
\begin{aligned}
t\|x(t)\| \leqq\; & t\|U(t,s)x(s)\| + t\int_s^t \|U(t,\tau)\|\,\|f(\tau,0)\|\,d\tau \\
& + \int_s^t K(t-\tau)\|U(t,\tau)\|\,\|x(\tau)\|\,d\tau \\
& + \int_s^t K\|U(t,\tau)\|\cdot\tau\|x(\tau)\|\,d\tau, \qquad 0<s<t<+\infty.
\end{aligned}
\tag{4.38}
$$

The first, the second, and the third terms of (4.38) are bounded for $t(>s)$. In the same way as Theorem 4.4, (4.38) implies that $t\|x(t)\|$ is bounded for $t>0$. Therefore we find inductively that $t^n\|x(t)\|$ is bounded for $t>0$. Next consider the inequality

$$
\begin{aligned}
t^{n+\beta}\|x(t)\| \leqq\; & t^{n+\beta}\|U(t,s)x(s)\| + t^{n+\beta}\int_s^t \|U(t,\tau)\|\,\|f(\tau,0)\|\,d\tau \\
& + \int_s^t K\|U(t,\tau)\| \sum_{j=0}^{n} {}_nC_j(t-\tau)^{n-j+\beta}\tau^j\|x(\tau)\|\,d\tau \\
& + \int_s^t K\|U(t,\tau)\| \sum_{j=0}^{n-1} {}_nC_j(t-\tau)^{n-j}\tau^{j+\beta}\|x(\tau)\|\,d\tau \\
& + \int_s^t K\|U(t,\tau)\|\,\tau^{n+\beta}\|x(\tau)\|\,d\tau.
\end{aligned}
\tag{4.39}
$$

In the same way as Theorem 4.4, (4.39) implies that the estimate (4.34) holds when $t\to\infty$. In the case where $n=0$, we also obtain the same conclusion.

(iii) Consider the inequality

$$
\begin{aligned}
& |\ln t| \int_s^t \|U(t,\tau)\|\,\|f(\tau,0)\|\,d\tau \\
& \leqq \int_s^t \{|\ln(t-\tau)| + |\ln\tau| + \ln 2\}\|U(t,\tau)\|\,\|f(\tau,0)\|\,d\tau.
\end{aligned}
\tag{4.40}
$$

Here we have used the inequality

$$
|\ln(x+y)| \leqq |\ln x| + |\ln y| + \ln 2, \qquad x>0, \qquad y>0.
$$

It follows from (4.3) and (4.35) that the right-hand side of (4.40) is bounded for $t>s$. Therefore, in the same way as Theorem 4.4, the estimate (4.36) follows from the inequality

$$
\begin{aligned}
|\ln t|\,\|x(t)\| \leqq\; & |\ln t|\,\|U(t,s)x(s)\| + |\ln t| \int_s^t \|U(t,\tau)\|\,\|f(\tau,0)\|\,d\tau \\
& + \int_s^t \{|\ln(t-\tau)| + |\ln\tau| + \ln 2\}K\|U(t,\tau)\|\,\|x(\tau)\|\,d\tau. \qquad \text{Q.E.D.}
\end{aligned}
$$

5. Asymptotic behavior II. In this section we assume that the conditions (1.2), (2.5), (4.1), (4.2), and (1.9) for $f(t,x)$ are satisfied. Furthermore, for the simplicity we assume that $k(c)$ satisfies

$$
k(c) \leqq Kc^a, \qquad c>0,
\tag{5.1}
$$

where $K>0$ and $a>0$ are some constants.

Consider (4.14) with $\omega=0$ and $s=0$. If $L_1L_2M_2\lambda^{\alpha-1}\Gamma(1-\alpha)<1$, then it follows from (4.13) that

$$\int_0^t \|U(t,\tau)\|\,d\tau \leqq \frac{M_1\lambda^{-(2-1/\beta)}\Gamma(2-1/\beta)}{1-L_1L_2M_2\lambda^{\alpha-1}\Gamma(1-\alpha)}, \qquad t>0.$$

If $L_1L_2M_2\lambda^{\alpha-1}\Gamma(1-\alpha)\geqq 1$, choose $t_0>0$ such that $L_1(t_0)L_2M_2\lambda^{\alpha-1}\Gamma(1-\alpha)<1$. In the same way as (4.12) we obtain

$$g(t,0)\leqq L_1L_2M_2\lambda^{\alpha-1}\Gamma(1-\alpha)+\int_0^{t_0}\|R_1(t,\sigma)\|g(\sigma,0)\,d\sigma$$
$$+\int_{t_0}^t \|R_1(t,\sigma)\|g(\sigma,0)\,d\sigma, \qquad t>t_0.$$

From the above inequality we easily obtain

$$\sup_{0<t<+\infty} g(t,0)\leqq \frac{L_1L_2M_2\lambda^{\alpha-1}\Gamma(1-\alpha)[1+\sup_{0\leqq t\leqq t_0} g(t,0)]}{1-L_1(t_0)L_2M_2\lambda^{\alpha-1}\Gamma(1-\alpha)}.$$

$\sup_{0\leqq t\leqq t_0} g(t,0)$ is easily estimated. Therefore we can estimate

$$\sup_{0<t<+\infty}\int_0^t \|U(t,\tau)\|\,d\tau=p \tag{5.2}$$

in a concrete form. Next we estimate $\|U(t,0)A^{-1}(0)\|$ for $0<t<+\infty$. We note the inequalities

$$\|U(t,0)A^{-1}(0)\|\leqq\|e^{tA(0)}A^{-1}(0)\|+\int_0^t \|e^{(t-\tau)A(\tau)}\|\,\|R(\tau,0)A^{-1}(0)\|\,d\tau,$$

$$\|R(t,0)A^{-1}(0)\|\leqq\|R_1(t,0)A^{-1}(0)\|+\int_0^t \|R_1(t,\tau)\|\,\|R(\tau,0)A^{-1}(0)\|\,d\tau,$$

$$\|e^{tA(0)}A^{-1}(0)\|\leqq\|A^{-1}(0)\|+\int_0^t \|e^{\tau A(0)}\|\,d\tau.$$

Similarly we can also estimate

$$\sup_{0<t<+\infty}\|U(t,0)A^{-1}(0)\|=q \tag{5.3}$$

in a concrete form by using the above inequalities.

If $\|f(t,0)\|$ is small, we obtain the following lemma:

LEMMA 5.1. *Suppose that* $f(t,x)$ *satisfies* (1.9) *and* (5.1) *and that the inequality*

$$r=\frac{a}{a+1}[(a+1)Kp]^{-1/a}-p\sup_{0\leqq t}\|f(t,0)\|>0 \tag{5.4}$$

holds. Then the global solution $x(t)$ *of* (3.1) *uniquely exists for* $x_0\in D(A)$ *satisfying*

$$\|A(0)x_0\|\leqq rq^{-1}, \tag{5.5}$$

and satisfies the estimate

$$\|x(t)\|\leqq[(a+1)Kp]^{-1/a}, \qquad 0\leqq t<+\infty.$$

Proof. In the successive approximations (3.2), let

$$e_n = \sup_{0 \leqq t} \|x_n(t)\|.$$

Since $x_n(t)$ satisfy the inequalities

$$\|x_{n+1}(t)\| \leqq \|U(t, 0)x_0\| + \int_0^t \|U(t, \tau)\| \{k(e_n)\|x_n(\tau)\| + \|f(\tau, 0)\|] \, d\tau,$$

it follows from (5.4) and (5.5) that the estimates

$$\text{(5.6)} \qquad \begin{aligned} &e_{n+1} \leqq \frac{a}{a+1}[(a+1)Kp]^{-1/a} + Kpe_n^{a+1}, \qquad n = 0, 1, \cdots, \\ &e_0 < [(a+1)Kp]^{-1/a} \end{aligned}$$

hold. From (5.6) we obtain

$$e_n < [(a+1)Kp]^{-1/a}, \qquad n = 0, 1, \cdots.$$

Therefore $x_n(t)$ converge to $x(t)$ uniformly on any finite closed interval in $[0, \infty)$. Clearly $x(t)$ defined in (3.3) is the unique solution of (3.1). Q.E.D.

Remark. If $f(t, x)$ satisfies (3.4) in addition to the assumptions of Lemma 5.1, $x(t)$ satisfies (1.1). If $f(t, x)$ satisfies (3.13) to (3.16), the equation (3.17) holds.

THEOREM 5.2. *Suppose that the assumptions of Lemma* 5.1 *are satisfied. Then we obtain*

$$\text{(5.7)} \qquad \overline{\lim_{t\to\infty}} \|x(t)\| \leqq \frac{(a+1)p}{a} \overline{\lim_{t\to\infty}} \|f(t, 0)\|,$$

where $x(t)$ is the solution of (3.1) *with $x_0 \in D(A)$ satisfying* (5.5).

Proof. For any $\varepsilon > 0$, choose $s > 0$ so that the estimates

$$\|x(t)\| \leqq \overline{\lim_{t\to\infty}} \|x(t)\| + \varepsilon, \qquad \|f(t, 0)\| \leqq \overline{\lim_{t\to\infty}} \|f(t, 0)\| + \varepsilon$$

hold for $t \geqq s$. Then, for $t > s$ we have

$$\begin{aligned} \|x(t)\| &\leqq \|U(t, s)x(s)\| + \int_s^t \|U(t, \tau)\| \left\{ \frac{\|x(\tau)\|}{(a+1)p} + \|f(\tau, 0)\| \right\} d\tau \\ &\leqq \|U(t, s)\| \, \|x(s)\| + \frac{1}{a+1} \left\{ \overline{\lim_{t\to\infty}} \|x(t)\| + \varepsilon \right\} + p \left\{ \overline{\lim_{t\to\infty}} \|f(t, 0)\| + \varepsilon \right\}. \end{aligned}$$

The above inequality and (4.3) imply that the estimate (5.7) holds. Q.E.D.

Remark. Under the assumptions of Theorem 5.2, we can derive the similar results to Theorem 4.5.

6. Asymptotic behavior III. In this section it is assumed that (1.2), (2.5), and (4.2) are satisfied and that $f(t, x)$ satisfies (3.13) to (3.16) and either (1.8) or (1.9). Furthermore it is assumed that the following conditions are satisfied:

(6.1) (i) $A(t)$, $t \geqq 0$, is strongly continuously differentiable on $D(A)$.

(6.2) (ii) $\displaystyle\int_0^\infty \|A'(t)A^{-1}(0)\| \, dt < +\infty, \quad \lim_{t\to\infty} \|A'(t)A^{-1}(0)\| = 0.$

(6.3) (iii) $\displaystyle\sup_{0 \leqq t} \|f(t, 0)\| < +\infty.$

From (6.2) we find that there exists a bounded linear operator B such that $A(t)A^{-1}(0)$ converges to B in norm as $t\to\infty$. Since the equation

$$B=A(t)A^{-1}(0)[I-A(0)A^{-1}(t)(A(t)A^{-1}(0)-B)]$$

holds for $t\geqq 0$, it follows that the bounded operator B^{-1} exists. Let $A(\infty)=BA(0)$ and let $D(A(\infty))=D(A)$ respectively. Then, $A(\infty)$ is a closed operator. Noting the inequality

$$\|A(t)A^{-1}(\tau)-I\|\leqq\|(A(t)-A(\tau))A^{-1}(0)\|\,\|A(0)A^{-1}(\tau)\|,\qquad t,\tau\geqq 0,$$

we find that $L_2(s)$ converges to 1 as $s\to\infty$.

We start with the following lemma:

LEMMA 6.1. *Suppose that there exists the unique solution $x(t)$ of* (1.1) *with* $x(0)=x_0$ *and that $x(t)$ is bounded on* $[0,\infty)$. *If $x(t)$ satisfies the estimate*

$$m=M_1\lambda^{-(2-1/\beta)}\Gamma(2-1/\beta)\varlimsup_{t\to\infty}\|Df(t,x(t))+\bar{D}f(t,x(t))\|_{L(E,E)}<1,\tag{6.4}$$

then we obtain

$$\varlimsup_{t\to\infty}\left\|\frac{d}{dt}x(t)\right\|\leqq(1-m)^{-1}M_1\lambda^{-(2-1/\beta)}\Gamma(2-1/\beta)\varlimsup_{t\to\infty}\|f_t(t,x(t))\|.\tag{6.5}$$

Remark 1. Let A (resp. B) be a linear (resp. antilinear) bounded operator. Then $\|A+B\|_{L(E,E)}$ is understood to be

$$\sup_{\|x\|\leqq 1}\|Ax+Bx\|.$$

Remark 2. Suppose that the estimates

$$\|Df(t,x)-Df(t,y)\|_{L(E,E)}\leqq k_1(c)\|x-y\|,$$
$$\|\bar{D}f(t,x)-\bar{D}f(t,y)\|_{L(E,E)}\leqq k_2(c)\|x-y\|$$

hold for t, x, and y satisfying $t\geqq 0$, $\|x\|\leqq c$, $\|y\|\leqq c$, where $k_1(c)$ and $k_2(c)$ are monotone nondecreasing functions of $c>0$ which are right continuous. Then the condition (6.4) can be written in the more concrete form by combining the above estimates with (4.27) or (5.7).

Proof. (6.5) is clear in the case where $\varlimsup_{t\to\infty}\|f_t(t,x(t))\|=+\infty$. Take sufficiently small $\varepsilon>0$. Then, from (3.17), (4.16), (6.2), and (6.3), there exists $s(\varepsilon)>0$ such that the inequalities

$$\begin{aligned}\left\|\frac{d}{dt}x(t)\right\|&\leqq\left\|V(t,s)\frac{d}{ds}x(s)\right\|+\int_s^t\|V(t,\sigma)\|\,\|A'(\sigma)A^{-1}(\sigma)f(\sigma,x(\sigma))\|\,d\sigma\\&\quad+\int_s^t\|V(t,\sigma)\|\,\|f_\sigma(\sigma,x(\sigma))\|\,d\sigma\\&\quad+\int_s^t\|V(t,\sigma)\|\,\|Df(\sigma,x(\sigma))+\bar{D}f(\sigma,x(\sigma))\|_{L(E,E)}\left\|\frac{d}{d\sigma}x(\sigma)\right\|\,d\sigma\\&\leqq\left\|V(t,s)\frac{d}{ds}x(s)\right\|+\varepsilon+\left\{M_1\lambda^{-(2-1/\beta)}\Gamma(2-1/\beta)\varlimsup_{t\to\infty}\|f_t(t,x(t))\|+\varepsilon\right\}\\&\quad+\int_s^t\|V(t,\sigma)\|\,\|Df(\sigma,x(\sigma))+\bar{D}f(\sigma,x(\sigma))\|_{L(E,E)}\left\|\frac{d}{d\sigma}x(\sigma)\right\|\,d\sigma,\end{aligned}\tag{6.6}$$

$$\int_s^t \|V(t,\sigma)\| \, \|Df(\sigma, x(\sigma)) + \bar{D}f(\sigma, x(\sigma))\|_{L(E,E)} \, d\sigma \leqq m + \varepsilon < 1, \tag{6.7}$$
$$s(\varepsilon) \leqq s < t < +\infty$$

hold. Therefore, in the same way as in Theorem 4.4, (6.7) implies (6.5). Q.E.D.

Next suppose that $f(t, x)$ satisfies the following additional condition:

For each $c>0$, there exists a bounded measurable function $k(t, c)$ of t which is integrable on $(0, \infty)$ and $k(t, c)$ satisfies

$$\|f_t(t,x)\| \leqq k(t,c), \quad \|x\| \leqq c, \quad t>0, \tag{6.8}$$
$$\lim_{t\to\infty} k(t,c) = 0.$$

From (6.8) we find that there exists a continuous function $f_\infty(x)$ such that $f(t, x)$ converges to $f_\infty(x)$ as $t \to \infty$ uniformly on each bounded set in E. Then we obtain the following assertion:

THEOREM 6.2. *Suppose that the unique solution $x(t)$ of* (1.1) *with $x(0) = x_0$ exists and that $f(t, x)$ satisfies* (6.8). *If $x(t)$ is bounded on $[0, \infty)$ and satisfies* (6.4), *then there exists $x(\infty) \in D(A)$ which satisfies the equation*

$$A(\infty)x(\infty) + f_\infty(x(\infty)) = 0, \tag{6.9}$$

and $x(t)$ converges to $x(\infty)$ as $t \to \infty$.

Proof. Let $\|x(t)\| \leqq c$, $t \geqq 0$. Then, it follows from (6.4) and (6.8) that

$$\lim_{t\to\infty} \left\| \frac{d}{dt} x(t) \right\| = 0. \tag{6.10}$$

By integrating the both sides of (6.6) with respect to t from s to T, we obtain the inequality

$$\int_s^T \left\| \frac{d}{dt} x(t) \right\| dt$$
$$\leqq \int_s^T \|V(t,s)\| \, dt \left\| \frac{d}{ds} x(s) \right\| + \int_s^T \|A'(\sigma)A^{-1}(\sigma) f(\sigma, x(\sigma))\| \, d\sigma \int_\sigma^T \|V(t,\sigma)\| \, dt$$
$$+ \int_s^T \|f_\sigma(\sigma, x(\sigma))\| \, d\sigma \int_\sigma^T \|V(t,\sigma)\| \, dt$$
$$+ \int_s^T \left\| \frac{d}{d\sigma} x(\sigma) \right\| d\sigma$$
$$\cdot \sup_{(s\leqq)\tau} \|Df(\tau, x(\tau)) + \bar{D}f(\tau, x(\tau))\|_{L(E,E)} \cdot \sup_{(s\leqq)\tau<T} \int_\tau^T \|V(t,\tau)\| \, dt,$$

where $T > s$ is arbitrary. Therefore the above inequality, (4.17), (6.2) to (6.4), and (6.8) imply that

$$\int_s^\infty \left\| \frac{d}{dt} x(t) \right\| dt < +\infty, \quad s > 0. \tag{6.11}$$

Consequently it follows that there exists $x(\infty) \in E$ such that

$$x(t) \to x(\infty), \quad f(t, x(t)) \to f_\infty(x(\infty)) \quad \text{as } t \to \infty. \tag{6.12}$$

Noting the inequality

$$\|A(\infty)A^{-1}(t)-I\| \leqq \|(A(\infty)-A(t))A^{-1}(0)\| \; \|A(0)A^{-1}(t)\|,$$

we obtain that $A(\infty)A^{-1}(t)$ converges to I in norm as $t \to \infty$ [11], [12]. Therefore it follows from (6.12) that

$$A(\infty)x(t)=A(\infty)A^{-1}(t)A(t)x(t) \to -f_\infty(x(\infty)), \quad t \to \infty.$$

Since $A(\infty)$ is closed, the above relation implies $x(\infty) \in D(A)$ and (6.9). Q.E.D.

Corresponding to the degree of decreasing of $k(t, c)$ and $\|A'(t)A^{-1}(0)\|$, the following assertion holds:

THEOREM 6.3. *Suppose that the assumptions of Theorem* 6.2 *are satisfied.*

(i) *If there exist* $d_1=d_1(c)>0$, $\delta_1=\delta_1(c)>0$, $d_2>0$, *and* $\delta_2>0$ *such that the estimates*

$$(6.13) \qquad \begin{aligned} k(t, c) &\leqq d_1 e^{-\delta_1 t}, \quad t>0, \\ \|A'(t)A^{-1}(0)\| &\leqq d_2 e^{-\delta_2 t}, \quad t>0, \end{aligned}$$

hold, then the solution $x(t)$ *of* (1.1) *satisfies the estimate*

$$(6.14) \qquad \|x(t)-x(\infty)\| \leqq d\, e^{-\delta t}, \quad t>0,$$

where $d>0$ *and* $\delta>0$ *are some constants.*

(ii) *If there exist* $d_1=d_1(c)>0$, $d_2>0$, *integers* $n_1=n_1(c) \geqq 1$, $n_2 \geqq 1$, $0<\beta_1=\beta_1(c) \leqq 1$, *and* $0<\beta_2 \leqq 1$ *such that the estimates*

$$(6.15) \qquad k(t, c) \leqq d_1 t^{-(n_1+\beta_1)}, \quad \|A'(t)A^{-1}(0)\| \leqq d_2 t^{-(n_2+\beta_2)}$$

hold when $t \to \infty$, *then* $x(t)$ *satisfies the estimate*

$$(6.16) \qquad \|x(t)-x(\infty)\| \leqq d\, t^{-(n+\beta-1)}$$

when $t \to \infty$, *where* $n+\beta=\min(n_1+\beta_1, n_2+\beta_2)$, *and d is some constant.*

Proof. (i) In the same way as Theorem 4.5 (i), we find that $e^{\nu t}\|(d/dt)x(t)\|$ is bounded when $t \to \infty$, where $\nu>0$ is some constant depending on δ_1, δ_2, and λ. Therefore (6.14) immediately follows from the equation

$$x(\infty)-x(t)=\int_t^\infty \frac{d}{d\tau}x(\tau)\, d\tau, \quad t>0.$$

The proof of (ii) is similar to the above arguments. Hence we omit it. Q.E.D.

If $A(t)=A$, $t \geqq 0$, we assume the condition $1/2<\beta<1$ weaker than (2.5) so that the semigroup e^{tA} is integrable on $(0, \infty)$. Then we have the following corollary:

COROLLARY 6.4. *Suppose that* $A(t)$ *and* $f(t, x)$ *are independent of t. If the solution* $x(t)$ *of* (1.1) *satisfies the estimate*

$$(6.17) \qquad M_1 \lambda^{-(2-1/\beta)}\Gamma(2-1/\beta) \varlimsup_{t \to \infty} \|Df(x(t))+\bar{D}f(x(t))\|_{L(E,E)}<1,$$

then there exists $x(\infty) \in D(A)$ *which satisfies the equation*

$$(6.18) \qquad Ax(\infty)+f(x(\infty))=0,$$

and $x(t)$ *converges to* $x(\infty)$ *exponentially.*

Proof. The proof is the same as that of Lemma 6.1 and Theorem 6.2.

Remark. In Corollary 6.4, it has not been assumed that $\|x(t)\|$ is bounded on $[0, \infty)$.

7. Examples. In this section we give some examples of linear partial differential equations whose coefficients $A(t)$ satisfy the estimate (1.2). It seems more important to give examples of (1.7) whose evolution operators have the integrable singularities rather than to give examples of (1.1), because there exist a great many kinds of nonlinear terms $f(t, x)$ satisfying (1.8) or (1.9).

In Example 1 through 6 we treat the case where $A(t)$ is independent of $t \geqq 0$.

Example 1. Consider the initial boundary value problem of the heat equation

$$\frac{\partial}{\partial t} u(t, x) = \frac{\partial^2}{\partial x^2} u(t, x), \quad t > 0, \quad 0 \leqq x \leqq 1, \tag{7.1}$$

$$u(0, x) = u_0(x), \quad 0 \leqq x \leqq 1, \tag{7.2}$$

$$\frac{\partial}{\partial x} u(t, 0) = \frac{\partial}{\partial x} u(t, 1) = 0. \tag{7.3}$$

Let $E = C^\alpha([0, 1])$, where $0 < \alpha < 1$. The norm in E is given by

$$\|u\|_\alpha = \sup_{0 \leqq x \leqq 1} |u(x)| + \sup_{0 \leqq x, y \leqq 1} \frac{|u(x) - u(y)|}{|x - y|^\alpha}. \tag{7.4}$$

The operator A and the domain $D(A)$ are given by

$$Au = \frac{d^2}{dx^2} u, \qquad D(A) = \left\{ u \in C^{2+\alpha}; \frac{d}{dx} u(0) = \frac{d}{dx} u(1) = 0 \right\}. \tag{7.5}$$

The resolvent $R(\mu) = (A - \mu I)^{-1}$ exists in the complex plane except for the non-positive real semiaxis and is represented as follows:

$$\begin{aligned} u(x) = R(\mu) f = &-\frac{\cosh \sqrt{\mu}\,(x-1)}{\sqrt{\mu} \sinh \sqrt{\mu}} \int_0^x \cosh \sqrt{\mu}\, \xi \cdot f(\xi)\, d\xi \\ &-\frac{\cosh \sqrt{\mu}\, x}{\sqrt{\mu} \sinh \sqrt{\mu}} \int_x^1 \cosh \sqrt{\mu}\, (\xi - 1) \cdot f(\xi)\, d\xi. \end{aligned} \tag{7.6}$$

Let any $\delta > 0$ and any ε, $0 < \varepsilon < \pi/2$, be given. In the following, we shall estimate $\|R(\mu)\|_\alpha$ on the sector Σ;

$$\Sigma = \left\{ \mu = \zeta + \delta;\ -\frac{\pi}{2} - \varepsilon \leqq \arg \zeta \leqq \frac{\pi}{2} + \varepsilon \right\}.$$

Let $\mu = re^{i\theta}$, $-\pi/2 - \varepsilon < \theta < \pi/2 + \varepsilon$, and $\beta = \cos(\theta/2)$. It is easy to see that the estimate

$$\|u\|_0 \leqq \frac{1}{r\beta} \|f\|_0 \tag{7.7}$$

holds, where $\|\cdot\|_0$ denotes the supremum norm of $C([0, 1])$.

Let $0 \leqq y < x \leqq 1$. Then from (7.6) we have

$$\begin{aligned}\frac{|u(x)-u(y)|}{|x-y|^{\alpha}} \leqq & \frac{1}{|x-y|^{\alpha}}\left|\frac{1}{\sqrt{\mu}\sinh\sqrt{\mu}}\int_y^x g(\xi; x, y)\{f(\xi)-f(y)\}\,d\xi\right| \\ & +\frac{1}{|x-y|^{\alpha}}\left|\frac{1}{\sqrt{\mu}\sinh\sqrt{\mu}}\int_y^x g(\xi; x, y)\,d\xi\cdot f(y)\right| \\ & +\frac{|\cosh\sqrt{\mu}\,(x-1)-\cosh\sqrt{\mu}\,(y-1)|}{|x-y|^{\alpha}|\sqrt{\mu}\sinh\sqrt{\mu}|}\left|\int_0^y \cosh\sqrt{\mu}\,\xi\cdot f(\xi)\,d\xi\right| \\ & +\frac{|\cosh\sqrt{\mu}\,x-\cosh\sqrt{\mu}\,y|}{|x-y|^{\alpha}|\sqrt{\mu}\sinh\sqrt{\mu}|}\left|\int_x^1 \cosh\sqrt{\mu}\,(\xi-1)\cdot f(\xi)\,d\xi\right|,\end{aligned} \tag{7.8}$$

where

$$g(\xi; x, y)=\cosh\sqrt{\mu}\,(x-1)\cdot\cosh\sqrt{\mu}\,\xi-\cosh\sqrt{\mu}\,y\cdot\cosh\sqrt{\mu}\,(\xi-1).$$

It is easy to see that the estimate

$$\text{the first term of (7.8)} \leqq \frac{2}{r\beta}\{\|f\|_{\alpha}-\|f\|_0\} \tag{7.9}$$

holds. Since the equation

$$\int_y^x g(\xi; x, y)\,d\xi=\frac{1}{4\sqrt{\mu}}\{(e^{\sqrt{\mu}(x-1)}-e^{\sqrt{\mu}(y-1)})(e^{\sqrt{\mu}x}-e^{\sqrt{\mu}y}) -(e^{\sqrt{\mu}(1-x)}-e^{\sqrt{\mu}(1-y)})(e^{-\sqrt{\mu}x}-e^{-\sqrt{\mu}y})\}$$

holds, we obtain the estimate

$$\begin{aligned}&\text{the second term of (7.8)} \\ &\leqq r^{\alpha/2-1}\frac{2\{2^{3/2-\alpha}|\sin(\theta/2)|^{\alpha}+(\cos(\theta/2))^{\alpha}\}e^{\sqrt{r\beta}}}{e^{\sqrt{r\beta}}-e^{-\sqrt{r\beta}}}\|f\|_0.\end{aligned} \tag{7.10}$$

For the third term of (7.8), we obtain

$$\text{the third term of (7.8)} \leqq \{2^{3/2-\alpha}|\sin(\theta/2)|^{\alpha}+(\cos(\theta/2))^{\alpha}\}\beta^{-1}r^{\alpha/2-1}\|f\|_0. \tag{7.11}$$

Similarly we obtain

$$\text{the fourth term of (7.8)} \leqq \{2^{3/2-\alpha}|\sin(\theta/2)|^{\alpha}+(\cos(\theta/2))^{\alpha}\}\beta^{-1}r^{\alpha/2-1}\|f\|_0. \tag{7.12}$$

Therefore (7.7) and (7.9) to (7.12) imply that the estimate

$$\|R(\mu)\|_{\alpha} \leqq \frac{M_{\Sigma}}{|\mu|^{1-\alpha/2}}, \qquad \mu\in\Sigma \tag{7.13}$$

holds, where $M_{\Sigma}>0$ is some constant. It is easy to see that (7.13) implies (1.2).

Remark. In this example we can shift the path of the integration in (1.5) from Γ_q to $\partial\Sigma$. Therefore the semigroup e^{tA}, $t>0$, of (7.1), (7.2), and (7.3) satisfies the estimates

$$\|A^n e^{tA}\| \leqq M_{n+1}e^{\gamma t}t^{-n-\alpha/2}, \qquad t>0, \quad n=0, 1, \cdots, \tag{7.14}$$

where $M_{n+1}>0$ and γ are some constants.

In the following Examples 2, 3, and 4, E is considered to be $C^{\alpha}([0, 1])$ and A is considered to be d^2/dx^2, respectively.

Example 2. Consider (7.1) and (7.2) under the boundary condition

$$u(t,0)=\frac{\partial}{\partial x}u(t,1)=0. \tag{7.15}$$

The domain $D(A)$ of A is given by

$$D(A)=\left\{u\in C^{2+\alpha};\ u(0)=\frac{d}{dx}u(1)=0\right\}. \tag{7.16}$$

The resolvent $R(\mu)=(A-\mu I)^{-1}$ exists in the same region as in Example 1 and is represented as

$$\begin{aligned} u(x)=R(\mu)f= &-\frac{\cosh\sqrt{\mu}\,(x-1)}{\sqrt{\mu}\cosh\sqrt{\mu}}\int_0^x \sinh\sqrt{\mu}\,\xi\cdot f(\xi)\,d\xi \\ &-\frac{\sinh\sqrt{\mu}\,x}{\sqrt{\mu}\cosh\sqrt{\mu}}\int_x^1 \cosh\sqrt{\mu}\,(\xi-1)\cdot f(\xi)\,d\xi. \end{aligned} \tag{7.17}$$

In a way similar to Example 1, $R(\mu)$ satisfies the estimate (7.13) on Σ.

Example 3. Consider (7.1) and (7.2) under the boundary condition

$$\frac{\partial}{\partial x}u(t,0)=u(t,1)=0. \tag{7.18}$$

The domain $D(A)$ of A is given by

$$D(A)=\left\{u\in C^{2+\alpha};\ \frac{d}{dx}u(0)=u(1)=0\right\}. \tag{7.19}$$

Then the resolvent $R(\mu)=(A-\mu I)^{-1}$ exists in the same region as in Example 1 and is represented as

$$\begin{aligned} u(x)=R(\mu)f= &\frac{\sinh\sqrt{\mu}\,(x-1)}{\sqrt{\mu}\cosh\sqrt{\mu}}\int_0^x \cosh\sqrt{\mu}\,\xi\cdot f(\xi)\,d\xi \\ &+\frac{\cosh\sqrt{\mu}\,x}{\sqrt{\mu}\cosh\sqrt{\mu}}\int_x^1 \sinh\sqrt{\mu}\,(\xi-1)\cdot f(\xi)\,d\xi. \end{aligned} \tag{7.20}$$

In a way similar to Example 1, $R(\mu)$ satisfies the estimate (7.13) on Σ.

Example 4. Consider (7.1) and (7.2) under the boundary condition

$$u(t,0)=u(t,1)=0. \tag{7.21}$$

The domain $D(A)$ of A is given by

$$D(A)=\{u\in C^{2+\alpha};\ u(0)=u(1)=0\}. \tag{7.22}$$

Then the resolvent $R(\mu)$ exists in the same region as in Example 1 and is represented as

$$\begin{aligned} u(x)=R(\mu)f= &\frac{\sinh\sqrt{\mu}\,(x-1)}{\sqrt{\mu}\sinh\sqrt{\mu}}\int_0^x \sinh\sqrt{\mu}\,\xi\cdot f(\xi)\,d\xi \\ &+\frac{\sinh\sqrt{\mu}\,x}{\sqrt{\mu}\sinh\sqrt{\mu}}\int_x^1 \sinh\sqrt{\mu}\,(\xi-1)\cdot f(\xi)\,d\xi. \end{aligned} \tag{7.23}$$

In a way similar to Example 1, $R(\mu)$ satisfies the estimate (7.13) on Σ.

Remark. For Example 4, more general results have been obtained. Consider the parabolic equation

$$\frac{\partial}{\partial t} u = \sum_{|\beta| \leqq 2m} a_\beta(t, x) \frac{\partial^{|\beta|}}{\partial^{\beta_1} x_1 \cdots \partial^{\beta_n} x_n} u, \qquad t>0, \quad x \in \bar{\Omega},$$

$$u(0, x) = u_0(x), \quad x \in \bar{\Omega},$$

$$\left.\frac{\partial^{|\beta|}}{\partial^{\beta_1} x_1 \cdots \partial^{\beta_n} x_n} u\right|_{\partial\Omega} = 0, \qquad |\beta| \leqq m-1,$$

where Ω is a domain in R^n with a sufficiently smooth boundary and the coefficients $a_\beta(t, x)$ are smooth in t and x. Let $E = C^\alpha(\bar{\Omega})$. W. von Wahl [13] obtained the estimate (7.13) in the case where Ω is bounded. In the case where Ω is unbounded, H. Kielhöfer [4], [5] obtained (7.13).

Example 5 ([6, p. 161]). Consider the following initial value problem of a system which is parabolic in the sense of Šilov:

$$\frac{\partial}{\partial t} v_1 = \frac{\partial^2}{\partial x^2} v_1,$$

$$\frac{\partial}{\partial t} v_2 = i \frac{\partial^3}{\partial x^3} v_1 + \frac{\partial^2}{\partial x^2} v_2, \qquad t>0, \quad x \in R^1, \tag{7.24}$$

$$v_1(0, x) = \phi_1(x), \qquad v_2(0, x) = \phi_2(x), \qquad x \in R^1.$$

Let $E = L^2(R^1)$. Then the semigroup e^{tA}, $t>0$, of (7.24) satisfies

$$\|e^{tA}\| \leqq M t^{-1/2}, \qquad t>0. \tag{7.25}$$

Example 6. Consider the following initial value problem of a system which is parabolic in the sense of Šilov:

$$\frac{\partial}{\partial t} v_1 = \Delta v_1,$$

$$\frac{\partial}{\partial t} v_2 = \left(i \frac{\partial^3}{\partial x_1^3} - \frac{\partial^2}{\partial x_2^2} \right) v_1 + \Delta v_2, \quad t>0, \quad x = (x_1, x_2) \in R^2, \tag{7.26}$$

$$v_1(0, x) = \phi_1(x), \quad v_2(0, x) = \phi_2(x), \qquad x \in R^2,$$

where Δ denotes the Laplacian in R^2. Let $E = L^2(R^2)$. By the Plancherel theorem, (7.26) is equivalent to the following system of ordinary differential equations in $L^2(R^2)$:

$$\frac{d}{dt} \tilde{v}_1 = -(p_1^2 + p_2^2) \tilde{v}_1,$$

$$\frac{d}{dt} \tilde{v}_2 = (p_1^3 + p_2^2) \tilde{v}_1 - (p_1^2 + p_2^2) \tilde{v}_2, \qquad t>0, \qquad p = (p_1, p_2) \in R^2, \tag{7.27}$$

$$\tilde{v}_1(0, p) = \tilde{\phi}_1(p), \qquad \tilde{v}_2(0, p) = \tilde{\phi}_2(p), \quad p \in R^2,$$

where $\tilde{\phi}(p)$ denotes the Fourier transform of $\phi(x) \in L^2(R^2)$. The semigroup e^{tA}, $t>0$, of (7.27) is the bounded operator of multiplication by the matrix $U(t; p)$

$$U(t; p) = \begin{pmatrix} e^{-(p_1^2+p_2^2)t} & 0 \\ t(p_1^3 + p_2^2)\, e^{-(p_1^2+p_2^2)t} & e^{-(p_1^2+p_2^2)t} \end{pmatrix}. \tag{7.28}$$

$\|e^{tA}\|$ is calculated according to the formula [6]

$$\|e^{tA}\| = \sup_{p\in R^2} \|U(t;p)\|_2, \qquad t>0, \tag{7.29}$$

where $\|U(t;p)\|_2$ is the norm of the matrix $U(t;p)$ as an operator in R^2. As is easily seen, the formula (7.29) implies that e^{tA}, $t>0$, satisfies the estimate (7.25).

Example 7. Consider the following initial boundary value problem of the heat equation:

$$\frac{\partial}{\partial t} u(t,x) = a(t)\frac{\partial^2}{\partial x^2} u(t,x) - b(t)u(t,x), \qquad t>0, \quad 0\leqq x\leqq 1, \tag{7.30}$$

$$u(0,x)=u_0(x), \qquad 0\leqq x\leqq 1, \tag{7.31}$$

$$\frac{\partial}{\partial x} u(t,0) = \frac{\partial}{\partial x} u(t,1) = 0, \tag{7.32}$$

where the coefficients $a(t)$ and $b(t)$ are Hölder continuous with exponent ρ and satisfy the inequalities

$$0<a_1\leqq a(t)\leqq a_2, \qquad 0<b_1\leqq b(t)\leqq b_2, \qquad t\geqq 0. \tag{7.33}$$

Let $E=C^\alpha([0,1])$. The operator $A(t)$, $t\geqq 0$, and the domain $D(A(t))$ are given by

$$A(t)=a(t)A-b(t)I, \quad D(A(t))=D(A), \qquad t\geqq 0, \tag{7.34}$$

where A and $D(A)$ have been defined in (7.5). Choose δ, $0<\delta<b_1/a_2$, in Example 1. Then, the resolvent $(A(t)-\mu I)^{-1}$ exists on the sector Σ_1;

$$\Sigma_1 = \left\{\mu = \zeta + \delta a_2 - b_1; -\frac{\pi}{2} - \varepsilon \leqq \arg \zeta \leqq \frac{\pi}{2} + \varepsilon\right\}.$$

$(A(t)-\mu I)^{-1}$ satisfies the estimate

$$\|(A(t)-\mu I)^{-1}\|_\alpha \leqq \frac{M_0}{a_1 \min(1, a_2^{\alpha/2-1})} \cdot \frac{1}{(1+|\mathrm{Im}\,\mu|)^{1-\alpha/2}}, \qquad \mu\in\Sigma_1.$$

Since $a(t)$ and $b(t)$ are Hölder continuous, the condition (1.3) is satisfied. It follows from (7.33) that the condition (4.2) is also satisfied. If $a(t)$ and $b(t)$ are continuously differentiable, then $A(t)$, $t\geqq 0$, is strongly continuously differentiable on $D(A)$. Furthermore, if $a'(t)$ and $b'(t)$ are integrable on $[0,\infty)$ and tend to 0 as $t\to\infty$, then the condition (6.2) is satisfied. Consider the conditions corresponding to (4.1) and (4.2). If we assume that the conditions

$$|a(t)-a(\tau)|\leqq N_1(s)(t-\tau)^\rho, \qquad s\leqq\tau\leqq t<+\infty,$$
$$|b(t)-b(\tau)|\leqq N_2(s)(t-\tau)^\rho, \qquad s\leqq\tau\leqq t<+\infty,$$

$$\lim_{s\to\infty} N_i(s)=0, \qquad i=1,2,$$

are satisfied in addition to (7.33), then (4.1) and (4.2) are satisfied. But these conditions do not guarantee the convergence of $A(t)A^{-1}(0)$ as $t\to\infty$. As such an example, we can consider $a(t)$ and $b(t)$ given by

$$a(t)=b(t)=\sin\sqrt{t+1}+2, \qquad t\geqq 0.$$

Remark. In this example, we note that the estimates similar to (7.14) hold for each $A(t)$. Therefore, for the existence and the uniqueness of the evolution operator $U(t, s)$, it is sufficient to assume the condition

$$0<\alpha<1, \quad \beta=1-\alpha/2, \quad 2(1-\beta)<\rho\leqq 1,$$

which is weaker than (2.5). The proof can be carried out in the same way as in § 2.

Acknowledgment. The author wishes to thank Professor Y. Sakawa for his encouragement and discussions, and Professor H. Tanabe for his many helpful suggestions.

REFERENCES

[1] A. FRIEDMAN, *Partial Differential Equations*, Holt, Rinehart and Winston, New York, 1969.
[2] E. HILLE AND R. S. PHILLIPS, *Functional Analysis and Semigroups*, American Mathematical Society, Providence, RI, 1957.
[3] T. KATO, *Nonlinear evolution equations in Banach spaces*, Proc. Symp. Appl. Math., 17 (1965), pp. 50–67.
[4] H. KIELHÖFER, *Halbgruppen und semilineare Anfangs-Randwertprobleme*, Manuscripta Math., 12 (1974), pp. 121–152.
[5] ———, *Existenz und Regularität von Lösungen semilinearer parabolischer Anfangs-Randwertprobleme*, Math. Z., 142 (1975), pp. 131–160.
[6] S. G. Krein, *Linear Differential Equations in Banach Space*, American Mathematical Society, Providence, RI, 1971.
[7] A. PAZY, *A class of semilinear equations of evolution*, Israel J. Math., 20 (1975), pp. 23–36.
[8] E. T. POULSEN, *Evolutionsgleichungen in Banach-Räumen*, Math. Z., 90 (1965), pp. 286–309.
[9] P. E. SOBOLEVSKII AND V. A. POGORELENKO, *Hyperbolic equations in Hilbert space*, Siberian Math. J., 8 (1967), pp. 123–145.
[10] H. TANABE, *On the equations of evolution in a Banach space*, Osaka J. Math., 12 (1960), pp. 363–376.
[11] ———, *Convergence to a stationary state of the solution of some kind of differential equations in a Banach space*, Proc. Japan Acad., 37 (1961), pp. 127–130.
[12] ———, *Evolution Equations*, Iwanami, Tokyo, 1975. (In Japanese.)
[13] W. VON WAHL, *Gebrochene Potenzen eines elliptischen Operators und parabolische Differentialgleichungen in Räumen hölderstetiger Funktionen*, Nachr. Akad. Wiss. Göttingen. Math.-Phys. Kl. II 1972, Nr. 11, pp. 231–258.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics

0036-1410/78/0904-0014 $01.00/0

NONLINEAR BOUNDARY VALUE PROBLEMS FOR ORDINARY DIFFERENTIAL EQUATIONS WITH A SMALL PARAMETER*

R. KENT NAGLE†

Abstract. The author is concerned with the solution of nonlinear boundary value problems. In particular, he deals with the problem of resonance for a nonselfadjoint system of real ordinary differential equations with homogeneous linear boundary conditions. An alternative method similar to the one used by J. K. Hale when studying periodic solutions is used to reduce the problem to solving a system of q equations in p unknowns where p is the number of independent solutions to the associated linear boundary value problem and q is the number of independent solutions to the associated adjoint linear boundary value problem. Conditions are given for solving the system of q equations in p unknowns by the implicit function theorem. Several examples are included to illustrate the method.

1. Introduction. We are concerned in this paper with the solution of nonlinear boundary value problems for a system of n ordinary differential equations with homogeneous linear boundary conditions of the form

$$x' = A(t)x + \varepsilon f(t, x, x', \varepsilon), \qquad t \in [a, b], \tag{1.1}$$

$$B_1 x(a) + B_2 x(b) = 0, \tag{1.2}$$

where $x = \text{col}(x_1, \cdots, x_n)$, ε is a small real parameter, $f = \text{col}(f_1, \cdots, f_n)$, $A(t)$ is an $n \times n$ matrix, and B_1, B_2 are constant $m \times n$ matrices. Mainly we are concerned with nonselfadjoint problems, in the sense that the underlying linear problem is nonselfadjoint. We shall be particularly interested in the difficult problems "at resonance," in the sense that the underlying linear homogeneous problem has nontrivial solutions.

In § 2 we discuss the assumptions made on the column vector $f(t, x, x', \varepsilon)$ and the matrices $A(t)$, B_1, and B_2. Also included in this section is the necessary background material concerning linear boundary value problems.

In §§ 3 and 4 we describe the method of the Cesari–Hale alternative type (Cesari [1], Hale [5], [6]) which we shall use throughout this paper. Namely we extend to homogeneous boundary conditions the alternative scheme considered by Hale [6, p. 262] for handling periodic solutions of nonselfadjoint problems. We have also allowed the nonlinear term to include the derivative x'. In § 3 we give the definitions of the projection operators P and Q and the definition of the partial inverse operator K. We show that these operators have the same properties as Hale's operators by the same name. In § 4 we use these operators to decompose (1.1) into the equivalent system of two equations

$$x = Px + \varepsilon K(I - Q) f(t, x, x', \varepsilon) \tag{1.3}$$

$$Qf(t, x, x', \varepsilon) = 0, \tag{1.4}$$

the auxiliary and bifurcation equations respectively.

In § 5 we show that the auxiliary equation (1.3) is solvable for ε small by an application of Banach's fixed point theorem for contraction maps. Thus system (1.3)–(1.4) is reduced to the bifurcation equation (1.4), which takes the form of a system of q equations in p unknowns (the alternative problem). Here p is the number of independent solutions to the associated linear boundary value problem, and q is the

* Received by the editors November 26 1975, and in final revised form July 13, 1977.

† Department of mathematics, University of South Florida, Tampa, Florida 33620. This work was done while the author was at the University of Michigan–Dearborn, Dearborn, Michigan.

number of independent solutions to the associated adjoint linear boundary value problem. For $q \leqq p$ and f smooth, it is enough to verify that the relevant $q \times p$ Jacobian matrix for $\varepsilon = 0$ has maximum rank q so that, by the implicit function theorem, system (1.1)–(1.2) has at least one solution for each small ε. The results of § 5 when used in finding period solutions to a system of real ordinary differential equations yields the same bifurcation equations (see equation (5.1)) as obtained by Gambill and Hale [4] and Hale [5] which had been obtained by a similar version of the alternative method for selfadjoint problems. Mawhin [7], using alternative methods and degree theory, considered similar systems of real ordinary differential equations with periodic boundary conditions.

Finally, in § 6, we illustrate the method developed in this paper by considering several examples.

2. Preliminaries and basic assumptions. We are interested in the nonlinear boundary value problem

$$x' = A(t)x + \varepsilon f(t, x, x', \varepsilon), \qquad t \in [a, b] \tag{2.1}$$

where $x' = dx/dt$, $x = \operatorname{col}(x_1, \cdots, x_n)$ is the unknown vector function of t, ε is a small real parameter, $A(t)$ is a given $n \times n$ matrix whose enteries are bounded measurable functions, and $f = \operatorname{col}(f_1, \cdots, f_n)$ is an $n \times 1$ vector function defined on $[a, b] \times R^{2n+1}$ whose entries are measurable in t for every (x, x', ε) and continuous in (x, x', ε) for every t. Moreover, we assume that for each pair of constants R_1 and R_2, there exist constants M and L such that whenever $|x|, |y| \leqq R_1$, $|x'|, |y'| \leqq R_2$, then we have for all $t \in [a, b]$,

$$\begin{aligned} &|f(t, x, x', \varepsilon)| \leqq M, \\ &|f(t, x, x', \varepsilon) - f(t, y, y', \varepsilon)| \leqq L\{|x - y| + |x' - y'|\}. \end{aligned} \tag{2.2}$$

Here $|\cdot|$ denotes the Euclidean norm in R^n.

On system (2.1) we impose the boundary conditions

$$B_1 x(a) + B_2 x(b) = 0, \tag{2.3}$$

where B_1 and B_2 are constant $m \times n$ matrices such that $W = (B_1 B_2)$ is an $m \times 2n$ matrix of rank m.

Many of the results in §§ 4 and 5 are valid under weaker assumptions on f. For example if we replace equation (2.1) by the more general equation

$$x' = A(t)x + g(t, x, x', \varepsilon),$$

then Theorem 4.1 is true for any g which satisfies the conditions (2.2), and the continuity assumptions made on f. Except for Theorem 5.3, which involves the vector $H(\alpha, \varepsilon)$ (see (5.1)), the results of § 5 are true for $g(t, x, x', \varepsilon)$, which satisfies our continuity assumptions and the following assumptions: For each pair of constants R_1 and R_2, there exist $M(\varepsilon)$ and $L(\varepsilon)$, nonnegative continuous functions of ε, such that $M(0) = L(0) = 0$, and whenever $|x|, |y| \leqq R_1$, $|x'|, |y'| \leqq R_2$, then we have for all $t \in [a, b]$,

$$|f(t, x, x', \varepsilon)| \leqq M(\varepsilon),$$

$$|f(t, x, x', \varepsilon) - f(t, y, y', \varepsilon)| \leqq L(\varepsilon)\{|x - y| + |x' - y'|\}.$$

Condition (2.2) is just the special case in which $g = \varepsilon f$, $M(\varepsilon) = \varepsilon M$, and $L(\varepsilon) = \varepsilon L$.

In the present paper we shall use the following notation. Let $(\mathrm{AC}\,[a, b])^n$ denote the set of n vector functions y whose components are absolutely continuous functions on $[a, b]$ and such that $B_1y(a)+B_2y(b)=0$. For y in $(\mathrm{AC}[a, b])^n$ define $\|y\|_1$ by $\|y\|_1=\sup_{a\leq t\leq b}|y(t)|+(b-a)^{-1}\int_a^b|y'(s)|\,ds$. This defines a norm on $(\mathrm{AC}[a, b])^n$, and $(\mathrm{AC}(a, b])^n$ is a Banach space with this norm. Let $(L_1[a, b])^n$ denote the set of equivalence classes of n vector functions y whose components are Lebesgue integrable over $[a, b]$. For y in $(L_1[a, b])^n$ define $\|y\|_0$ by $\|y\|_0=(b-a)^{-1}\int_a^b|y(s)|\,ds$. This defines a norm on $(L_1[a, b])^n$ which makes $(L_1[a, b])^n$ a Banach space.

If T is a linear operator between two Banach spaces $(B_a, \|\cdot\|_a)$ and $(B_b, \|\cdot\|_b)$, the operator norm of T is given by $\|T\|$ where $\|T\|=\sup\{\|Tf\|_b: f\in B_a, \|f\|_a=1\}$.

The linear part of system (2.1), (2.3) is

$$x'=A(t)x, \qquad t\in[a, b], \tag{2.4}$$

where x satisfies the boundary conditions

$$B_1x(a)+B_2x(b)=0. \tag{2.5}$$

The adjoint boundary problem for (2.4)–(2.5) is

$$\tilde{y}'=-\tilde{y}A(t) \quad \text{or} \quad y'=-\tilde{A}(t)y, \tag{2.6}$$

with adjoint boundary conditions

$$\left.\begin{array}{l}\tilde{y}(a)=\tilde{\gamma}B_1\\ \tilde{y}(b)=-\tilde{\gamma}B_2\end{array}\right\} \quad \text{or} \quad \left\{\begin{array}{l}y(a)=\tilde{B}_1\gamma\\ y(b)=-\tilde{B}_2\gamma\end{array}\right. \tag{2.7}$$

where γ is an arbitrary vector in R^m and $\tilde{}$ denotes the transpose.

Associated with the boundary value problem (2.4)–(2.5) is the nonhomogeneous boundary problem

$$x=A(t)x+f(t), \qquad t\in[a, b] \tag{2.8}$$

with boundary conditions

$$B_1x(a)+B_2x(b)=0, \tag{2.9}$$

where $f=\mathrm{col}\,(f_1, \cdots, f_n)$ is assumed to be integrable in each component over $[a, b]$.

The following theorem, which is essentially known and referred to as the "Fredholm alternative," illustrates the connection between the nonhomogeneous boundary value problem and the adjoint boundary value problem. For a discussion of the properties of boundary value problems for differential equations, see Coddington and Levinson [2], and for differential systems, see Cole [3]. For a discussion of differential systems and a proof of the Fredholm alternative in this context, see Nagle [8].

THEOREM 2.1. *The nonhomogeneous boundary problem* (2.8)–(2.9) *has a solution if and only if*

$$\int_a^b z(s)f(s)\,ds=0$$

for every solution $z(t)$ to the adjoint boundary problem (2.6)–(2.7).

3. Definitions of the operators U, V, P, Q, and K. Let U be an $n\times p$ matrix whose p columns form a basis for the solutions to the boundary value problem (2.4)–(2.5), and let V be a $q\times n$ matrix whose q rows form a basis for the solutions to the adjoint boundary value problem (2.6)–(2.7). Let $c=\int_a^b\tilde{U}(s)U(s)\,ds$ and $d=\int_a^b V(s)\tilde{V}(s)\,ds$. The $p\times p$ matrix c and the $q\times q$ matrix d are nonsingular.

For y in $(\mathrm{AC}[a, b])^n$ we define the projection P: $(\mathrm{AC}[a, b])^n\to(\mathrm{AC}[a, b])^n$ by $Py(t)=U(t)c^{-1}\int_a^b\tilde{U}(s)y(s)\,ds$. It follows from the definition of U that the range of P is all of the subspace spanned by the solutions to (2.4)–(2.5).

In a similar fashion we define the projection $Q: (L_1[a, b])^n \to (L_1[a, b])^n$ by $Qg(t) = \tilde{V}(t)d^{-1} \int_a^b V(s)g(s)\, ds$. Similarly the range of Q is the subspace spanned by the solutions to the adjoint boundary problem (2.6)–(2.7). Straightforward calculations show that P and Q are bounded linear projections in their respective spaces.

In the next theorem the map K is defined from the kernel of Q into $(\mathrm{AC}\,[a, b])^n$. The map $K(I-Q)$ is shown to be a continuous linear transformation of $(L_1[a, b])^n$ into $(\mathrm{AC}\,[a, b])^n$.

THEOREM 3.1. *If h is in $(L_1[a, b])^n$, then a necessary and sufficient condition that the boundary value problem*

$$x' = A(t)x + h(t), \tag{3.1}$$

$$B_1x(a) + B_2x(b) = 0 \tag{3.2}$$

have a solution is that $Qh = 0$. If $Qh = 0$, then there exists a unique solution Kh of (3.1)–(3.2) *such that $PKh = 0$. Furthermore, $K(I-Q)$ is a continuous linear mapping of $(L_1[a, b])^n$ into $(AC\,[a, b])^n$.*

Proof. The first part of the theorem follows immediately from the definition of Q and the Fredholm alternative. For h in the range of Q, define $Kh = 0$. For h in the null space of Q we define Kh as follows.

Let $X(t)$ be a fundamental matrix for the homogeneous equation (2.4). If $X(a)x_0$ is the initial value for a solution $y(t)$ to (3.1)–(3.2), then by the variation of parameters formula we have $y(t) = X(t) \int_a^t X^{-1}(s)h(s)\, ds + X(t)x_0$. Since $y(t)$ must satisfy (3.2), it follows that $B_1X(a)x_0 + B_2X(b)x_0 = -B_2X(b) \int_a^b X^{-1}(s)h(s)\, ds$ or $Dx_0 = \beta$ where $D = B_1X(a) + B_2X(b)$ and $\beta = -B_2X(b) \int_a^b X^{-1}(s)h(s)\, ds$.

Let D^* be a partial right inverse for D; i.e. D^* maps the range of D into R^n in such a way that $DD^* = I$ on the range of D. To construct D^* let $\{v_1, \cdots, v_n\}$ be a basis for R^n. Since the set $\{Dv_1, \cdots, Dv_m\}$ spans the range of D, it may be reduced to a basis $\{Dv_1, \cdots, Dv_m\}$ with possibly a relabeling of the subscripts. Define $D^*(Dv_i) = v_i$ for $i = 1, \cdots, m$. Now extend D^* linearly to the entire range of D. Let D^* be zero off the range of D.

Since $Qh = 0$, β must be in the range of D, hence $D^*\beta$ corresponds to the initial value of a solution $y^*(h)$ of (3.1)–(3.2). We define K on the null space of Q by $Kh = (I-P)K_0(h) = (I-P)y^*(h)$. It follows that K is a linear mapping of $(L_1[a, b])^n$ into $(\mathrm{AC}\,[a, b])^n$. Moreover, it follows immediately from our definition of K that if $Qh = 0$, then Kh is a solution to (3.1)–(3.2) and $PKh = 0$.

Let $Qh = 0$. To show Kh is the unique solution to (3.1)–(3.2) such that $PKh = 0$, assume r is also a solution to (3.1)–(3.2) such that $Pr = 0$. Let $r^* = r - Kh$, then $(r^*)' = r' - (Kh)' = A(t)(r - Kh) = A(t)r^*$. Hence, r^* is a solution to the homogeneous boundary problem (2.4)–(2.5). Since r^* is in the range of the projection P, then $Pr^* = r^*$. So $PKh = P(r + r^*) = r^*$. Since $PKh = 0$, then we must have $r^* = 0$, and hence $r = Kh$. This proves the uniqueness.

It remains to prove that $K(I-Q)$ is bounded. Let $K_0h = x^*(h)$ for h in the null space of Q. Since $Kh = (I-P)K_0(h)$ for h in the null space of Q and since $I-P$ is bounded, it suffices to show that K_0 is a bounded mapping from the null space of Q into $(\mathrm{AC}\,[a, b])^n$. Let h be in the null space of Q. From the definition of K_0 it follows that

$$\begin{aligned} |(K_0h)(a)| &\leqq \|D^*\|\,\|B_2\|\,\|X(b)\|\,\|X^{-1}\|\,(b-a)\,\|h\|_0 \\ &= C_1\|h\|_0 \end{aligned} \tag{3.3}$$

where C_1 is a constant defined by (3.3), $\|D^*\|=\sum_{i,j}|d_{ij}^*|$, $\|B_2\|=\sum_{i,j}|b_{ij}|$, $\|X(b)\|=\sum_{i,j}|X(b)_{ij}|$, and $\|X^{-1}\|$ is a bound for the sum of the absolute values of the entries of X^{-1} in the closed interval $[a, b]$.

Since K_0h is a solution to (3.1), integrating we have, for $t\in[a, b]$, $|(K_0h)(t)|\leqq |(K_0h)(a)|+\int_a^t|h(s)|\,ds+\int_a^t|A(s)(K_0h)(s)|\,ds$. Hence, by the generalized Gronwall inequality (see Hale [6, p. 36]) it follows that

$$|(K_0h)(t)|\leqq|(K_0h)(a)|+\int_a^t|h(s)|\,ds$$

$$+\int_a^t|A(s)|\left\{|(K_0h)(a)|+\int_a^t|h(u)|\,du\right\}\exp\left\{\int_s^t|A(r)|\,dr\right\}ds \tag{3.4}$$

$$\leqq[C_1+(b-a)+M_A(b-a)\{C_1+(b-a)\}\exp\{(b-a)M_A\}]\|h\|_0.$$

Hence

$$\sup_{a\leqq t\leqq b}|(K_0h)(t)|\leqq C_2\|h\|_0, \tag{3.5}$$

where C_2 is the constant defined by equation (3.4).

Since K_0h satisfies equation (3.1), we have for $t\in[a, b]$,

$$|(K_0h)'(t)|=|A(t)(K_0h)(t)+h(t)|$$
$$\leqq M_AC_2\|h\|_0+|h(t)|.$$

Hence,

$$(b-a)^{-1}\int_a^b|(K_0h)'(t)|\,dt\leqq(M_AC_2+1)\|h\|_0. \tag{3.6}$$

It now follows from equations (3.5) and (3.6) that

$$\|(K_0h)\|_1\leqq(C_2+M_AC_2+1)\|h\|_0.$$

Hence, K_0 is a bounded map from the kernel of Q into $(\mathrm{AC}\,[a, b])^n$.

This completes the proof of Theorem 3.1.

COROLLARY 3.1. *If g is in $(L_1[a, b])^n$ and α is a given p vector, then the unique solution of the boundary value problem*

$$x'=A(t)x+(I-Q)g(t),\qquad B_1x(a)+B_2x(b)=0$$

with $Px=U(\cdot)\alpha$, is given by $x=U\alpha+K(I-Q)g$ and $x\in(AC\,[a, b])^n$.

Proof. Just differentiate $U\alpha+K(I-Q)g$.

4. The alternative scheme. In this section we split the boundary value problem (2.1), (2.3) into an equivalent system of two equations.

THEOREM 4.1. *For each fixed ε, the boundary value problem* (2.1), (2.3) *has a solution $x(t)$ if and only if $x(t)$ satisfies the boundary condition* (2.3) *and the system*

$$x=Px+\varepsilon K(I-Q)f(t, x, x', \varepsilon), \tag{4.1}$$

$$Qf(t, x, x', \varepsilon)=0. \tag{4.2}$$

Proof. Let x satisfy the boundary condition (2.3).

Let $w=x-Px$. Since P is idempotent, $Pw=0$. It follows from the definition of P that Px is a solution to $y'=A(t)y$ and satisfies the boundary condition (2.3). If x is a solution to (2.1), then x is a solution to the nonhomogeneous equation $y'=A(t)y+\varepsilon f(t, x, x', \varepsilon)$ and Theorem 3.1 implies that $Qf(t, x, x', \varepsilon)=0$. Now since x

satisfies the above nonhomogeneous equation and Px the homogeneous equation it follows that $w = x - Px$ satisfies the boundary value problem

$$y' = A(t)y + \varepsilon f(t, x, x', \varepsilon), \qquad B_1y(a) + B_2y(b) = 0. \tag{4.3}$$

By Corollary 3.1 we have $\varepsilon K(I-Q)f(t, x, x', \varepsilon)$ is the unique solution to (4.3) with $Py = 0$. But w also satisfies (4.3) with $Pw = 0$, hence $w = \varepsilon K(I-Q)f(t, x, x', \varepsilon)$.

Let x be a solution of the system (4.1), (4.2). Since $Qf(t, x, x', \varepsilon) = 0$, it follows from Corollary 3.1 that x is a solution to (2.1), (2.3). This completes the proof of Theorem 4.1.

Equation (4.1) is referred to as the auxiliary equation and equation (4.2) as the bifurcation equation.

5. Solving the auxiliary and bifurcation equations. The following theorem shows that the auxiliary equation always has a solution if ε is small.

THEOREM 5.1. *There exists $\rho > 0$ and $\varepsilon_0 > 0$ such that, for any constant p vector α, $|\alpha| \leqq \rho$ and ε such that $|\varepsilon| \leqq \varepsilon_0$, then there exists a unique n vector x^*, $x^* = x^*(\alpha, \varepsilon)$ such that*

$$x^* = U\alpha + \varepsilon K(I-Q)f(\cdot, x^*, x^{*\prime}, \varepsilon),$$

and x^* satisfies the boundary conditions (2.3). *Furthermore, if there is an $\alpha = \alpha(\varepsilon)$ with $|\alpha(\varepsilon)| \leqq \rho$ for $|\varepsilon| \leqq \varepsilon_0$, such that*

$$G(\alpha(\varepsilon), \varepsilon) \equiv Qf(\cdot, x^*(\alpha(\varepsilon), \varepsilon), x^{*\prime}(\alpha(\varepsilon), \varepsilon)) = 0,$$

then $x^*(\alpha(\varepsilon), \varepsilon)$ is a solution of the boundary value problem (2.1), (2.3).

Proof. The columns of U are continuous over $[a, b]$, hence bounded. Choose $\rho > 0$ so that $|\alpha| \leqq \rho$, α a p vector, implies $\|U\alpha\|_1 = \delta$ and $\delta < \min(R_1, R_2/B)$ where B is a bound for the sum of the absolute values of the entries of A.

Fix $\rho > 0$. Let α be a p vector with $|\alpha| \leqq \rho$. Define the subset S_α of $(\mathrm{AC}[a, b])^n$ by taking $S_\alpha = \{y \text{ in } (\mathrm{AC}[a, b])^n$: y satisfies (2.3), $Py = U\alpha$, $|y(t)| \leqq R_1$ for $t \in [a, b]$, and $|y'(t)| \leqq R_2$ for almost all $t \in [a, b]\}$. Let S be the union of the S_α for $|\alpha| \leqq \rho$, where α is a p vector. The set S is closed in $(\mathrm{AC}[a, b])^n$ as is S_α for each α. On S we define a family of maps $F(\alpha, \varepsilon)$ as follows. For $y \in S$, $F(\alpha, \varepsilon)y = U\alpha + \varepsilon K(I-Q)f(\cdot, y, y', \varepsilon)$. The maps $F(\alpha, \varepsilon)$ map S into $(\mathrm{AC}[a, b])^n$. It follows from the definitions of U and $K(I-Q)$ in § 3 that $F(\alpha, \varepsilon)y$ satisfies (2.3). For each $t \in [a, b]$

$$|(F(\alpha, \varepsilon)y)(t)| \leqq \delta + \varepsilon M\|K(I-Q)\|,$$

and for almost all $t \in [a, b]$

$$|(F(\alpha, \varepsilon)y)'(t)| \leqq B\delta + \varepsilon M\{\|K(I-Q)\| + 1 + M_v^2 M_d(b-a)\},$$

where M_v is a bound for the sum of entries of $V(t)$ and M_d is the sum of the absolute values of the entries of d^{-1}. Since $\delta < R_1$ and $B\delta < R_2$, we can choose ε_0 so that $F(\alpha, \varepsilon)$ maps S into S for $|\varepsilon| \leqq \varepsilon_0$. In fact since $P(F(\alpha, \varepsilon)y) = U\alpha$, $F(\alpha, \varepsilon)$ maps S_α into S_α.

For $y, z \in S$ we have

$$\|F(\alpha, \varepsilon)y - F(\alpha, \varepsilon)z\|_1 \leqq \varepsilon L\|K(I-Q)\| \, \|y - z\|_1,$$

so if we choose ε_0 such that $\varepsilon_0 L\|K(I-Q)\| < 1$, then the family $\{F(\alpha, \varepsilon): |\alpha| \leqq \rho, |\varepsilon| \leqq \varepsilon_0\}$ is a uniform family of contractions from S into S. Hence by the contraction mapping principle each $F(\alpha, \varepsilon)$ has a fixed point in S. In fact the fixed point lies in S_α.

We have shown that there exist $\rho>0$ and $\varepsilon_0>0$ such that for $|\alpha|\leqq\rho$, $|\varepsilon|\leqq\varepsilon_0$, then there is an x^* in S_α, such that $x^*=U\alpha+\varepsilon K(I-Q)f(\cdot, x^*, x^{*\prime}, \varepsilon)$.

If there is an $\alpha=\alpha(\varepsilon)$ with $|\alpha(\varepsilon)|\leqq\rho$ for $|\varepsilon|\leqq\varepsilon_0$, and $G(\alpha(\varepsilon), \varepsilon)=0$, then $x^*(\alpha(\varepsilon), \varepsilon)$ satisfies both the auxiliary and the bifurcation equations. Hence, it follows from Theorem 4.1 that x^* is a solution to (2.1), (2.3). This completes the proof of Theorem 5.1.

The solution to the auxiliary equation has several important properties which are a consequence of the Uniform Contraction Principle. We state the following theorem without proof (see Hale [6, p.7]).

THEOREM 5.2. (Uniform Contraction Principle). *If F is a closed subset of a Banach space X, G a subset of a Banach space Y, $T_y: F\to F$, y in G, is a uniform contraction on F (i.e. the T_y's are all contractions on F with the same constant) and T_yx is continuous in y for each fixed x in F, then the unique fixed point $g(y)$ of T_y, y in G, is continuous in y. Furthermore, if F, G are the closures of open sets F^0, G^0, and T_yx has continuous first derivatives in y when x is fixed and in x when y is fixed, then $g(y)$ has a continuous first derivative with respect to y in G^0.*

Our first result involves the uniqueness of the solution to the auxiliary equation.

COROLLARY 5.1. *Let S and $F(\alpha, \varepsilon)$ be as defined in the proof of Theorem* 5.1. *If for $|\varepsilon|\leqq\varepsilon_0$, $\bar{x}(\varepsilon)\in S$ (hence $P\bar{x}=U\alpha(\varepsilon)$ with $|\alpha(\varepsilon)|\leqq\rho$) and $\bar{x}(\varepsilon)$ is a solution to* (2.1), (2.3) *then $\bar{x}(\varepsilon)=x^*(\alpha(\varepsilon), \varepsilon)$ and $G(\alpha(\varepsilon), \varepsilon)=0$.*

Proof. The corollary follows from the uniqueness of the fixed point of the mapping $F(\alpha(\varepsilon), \varepsilon)$.

COROLLARY 5.2. *As defined in Theorem* 5.1 *let $x^*=x^*(\alpha, \varepsilon)$ be the solution to the auxiliary equation* (4.1) *for all $|\alpha|\leqq\rho$ and $|\varepsilon|\leqq\varepsilon_0$. Then x^* depends continuously on α and ε. Furthermore, if $f(\cdot, x, y, \varepsilon)$ has continuous partial derivatives with respect to x, y, and ε, then x^* is continuously differentiable with respect to α and ε.*

Proof. $F(\alpha, \varepsilon)$, as defined in the proof of Theorem 5.1, depends continuously on α and ε since f is a continuous function of ε. Hence by the Uniform Contraction Principle, the fixed point x^* depends continuously on α and ε. If $f(\cdot, x, y, \varepsilon)$ has a continuous derivative with respect to y when α and ε are fixed, and has a continuous derivative with respect to ε (or α) whenever y and α (or ε) are fixed, then x^* is continuously differentiable with respect to α and ε.

Corollary 5.2 says that $x^*(\alpha, \varepsilon)$ is Fréchet differentiable with respect to α when $f(\cdot, x, y, \varepsilon)$ is differentiable with respect to x, y, and ε. Since $x^*(\alpha, \varepsilon)$ is in $(\mathrm{AC}\,[a, b])^n$, $x^*(\alpha, \varepsilon)$ differentiable with respect to α implies $x^*(\alpha, \varepsilon)$ is differentiable with respect α as a vector function in $(L_1[a, b])^n$, and $x^{*\prime}(\alpha, \varepsilon)$ is differentiable with respect to α as a vector function in $(L_1(a, b])^n$.

As a consequence of the continuity of $x^*(\alpha, \varepsilon)$ with respect to ε, it follows that $x^*(\alpha, 0)=U\alpha$.

Theorem 5.1 assures us that for small ε the auxiliary equation always has a solution. Thus the boundary value problem (2.1), (2.3) is reduced to solving $Qf(t, x^*, x^{*\prime}, \varepsilon)=0$ where $x^*(\alpha, \varepsilon)$ is the solution to the auxiliary equation (4.1) for a particular α and ε. From the definition of Q it follows that $Qf(t, x^*, x^{*\prime}, \varepsilon)=0$ if and only if $\int_a^b V(s)f(s, x^*(s), x^{*\prime}(s), \varepsilon)\,ds=0$. Let

$$H(\alpha, \varepsilon)=\int_a^b V(s)f(s, x^*(s), x^{*\prime}(s), \varepsilon)\,ds. \tag{5.1}$$

The boundary value problem (2.1), (2.3) has a solution if and only if $H(\alpha, \varepsilon)=0$ has a solution for $|\varepsilon|\leqq\varepsilon_0$ and $|\alpha|\leqq\rho$. The equation $H(\alpha, \varepsilon)=0$ is also referred to as the *bifurcation equation* as is equation (4.2).

From the definition of V it follows that $H(\alpha, \varepsilon)$ is a system of q equations where q is the number of linearly independent solutions to the adjoint boundary value problem (2.6)–(2.7). As defined in Theorem 5.1, α is a constant p vector where p is the number of linearly independent solutions to the boundary value problem (2.4)–(2.5). For each small ε, to solve the bifurcation equation (5.1) we must solve a system of q equations in p unknowns. This can often be done using the implicit function theorem.

THEOREM 5.3. *Suppose ρ, ε_0, and $x^*(\alpha, \varepsilon)$ are defined as in Theorem* 5.1 *and $H(\alpha, \varepsilon)$, a system of q equations in $p+1$ unknowns, is defined as in equation* (5.1). *Assume $p \geqq q$ and $H(\alpha, \varepsilon)$ is continuously differentiable with respect to both α and ε for all $|\alpha| \leqq \rho$ and $|\varepsilon| \leqq \varepsilon_0$. If there is a p vector $\bar{\alpha}$ such that $|\bar{\alpha}| < \rho$, $H(\bar{\alpha}, 0) = 0$, and $H(\bar{\alpha}, 0)/\partial\alpha$ has rank q, then there is an $\varepsilon_1 > 0$ and a solution $x^*(\alpha(\varepsilon), \varepsilon)$, $|\varepsilon| \leqq \varepsilon_1$, of the boundary value problem* (2.1), (2.3).

Proof. The hypotheses on $H(\alpha, \varepsilon)$ and the implicit function theorem imply there is an ε_1, $0 \leqq \varepsilon_1 \leqq \varepsilon_0$, such that equation (5.1) has solution $\alpha(\varepsilon)$, $|\alpha(\varepsilon)| \leqq \rho$, for all ε, $|\varepsilon| \leqq \varepsilon_1$. But this implies $x^*(\alpha(\varepsilon), \varepsilon)$ is a solution to the boundary value problem (2.1), (2.3). This completes the proof of the theorem.

If $f(t, x, y, \varepsilon)$ is continuously differentiable with respect to x, y, and ε and there exists continuous functions $W(t)$ and $Z(t)$ so that for $|x| \leqq R_1$, $|y| \leqq R_2$, $|\varepsilon| < \varepsilon_0$,

$$|f_x(t, x, y, \varepsilon)| \leqq W(t),$$

and

$$|f_y(t, x, y, \varepsilon)| \leqq Z(t),$$

for all $t \in [a, b]$, then $H(\alpha, \varepsilon)$ is continuously differentiable with respect to α and ε.

Once one knows that the boundary value problem (2.1), (2.3) has a solution for fixed α and ε, $|\alpha| \leqq \rho$ and $|\varepsilon| \leqq \varepsilon_0$, then the solution to (2.1), (2.3) may be found by the method of successive approximation. One may begin the approximation with an element $y_0 = U\alpha = x^*(\alpha, 0)$; then the sequence $\{y_n\}$ defined by $y_{n+1} = F(\alpha, \varepsilon)y_n$, $n = 0, 1, 2, \cdots$ converges in $(\mathrm{AC}[a, b])^n$ to $x^*(\alpha, \varepsilon)$, the solution to (2.1), (2.3). Moreover, y_n will contain only terms up to order n in ε.

6. Applications and examples. In this section we consider specific problems which can be solved using the methods and results of the previous sections. Our emphasis is on the resonance cases which are harder to handle.

Example 1.

$$x'' = \varepsilon\{-e^{x''} + x^2 \sin^2 t\}. \tag{6.1}$$

We are interested in 2π-periodic solutions. Here, $p = 1$ and $q = 1$. Equation (6.1) may be written as the system

$$y' = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} y + \varepsilon g(t, y, y'),$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} y(0) + \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} y(2\pi) = 0,$$

where $y = \mathrm{col}\,(y_1, y_2)$, $y_1 = x$, and $y_2 = x'$ and $g(t, y, y') = \mathrm{col}\,(0, -e^{x''} + x^2 \sin^2 t)$. For this system $U(t) = \mathrm{col}\,(1, 0)$ and $V(t) = (0, 1)$. The auxiliary equation has a solution of the form $y(t) = U\alpha + \varepsilon K(I - Q)f(t, y, y', \varepsilon)$ where α is a constant.

Since $U = \mathrm{col}\,(1, 0)$ and $y_1 = x$, the solution to the auxiliary equation may be written in the form $y_1(t) = x(t) = \alpha + O(\varepsilon)$ where α is a constant. To find an α and ε such that the conditions of Theorem 5.3 are satisfied we calculate $H(\alpha, \varepsilon)$ from equation (5.1). Hence

$$H(\alpha, \varepsilon)=\int_0^{2\pi} V(t)g(t, y(t), y'(t))\,dt$$

$$=\int_0^{2\pi} -e^{O(\varepsilon)}+(\alpha^2+O(\varepsilon))\sin^2 t\,dt.$$

So $H(\alpha, 0)=\pi\alpha^2-2\pi$ and $\partial H(\alpha, 0)/\partial\alpha=2\pi\alpha$. It now follows from Theorem 5.3 that for $\alpha=\pm\sqrt{2}$ and ε small we have a 2π-periodic solution to equation (6.1) of the form $x(t)=\pm\sqrt{2}+O(\varepsilon)$.

Example 2.

$$x''+\sigma^2 x=\varepsilon\{(1-x^2)x'+a\sigma^{-1}x''+b\sigma(\cos(\sigma t+\alpha))\}. \tag{6.2}$$

We are interested in $2\pi/\sigma$-periodic solutions. Here, $p=2$ and $q=2$. Equation (6.2) can be written as the system

$$y'=\begin{pmatrix}0 & 1\\ -\sigma^2 & 0\end{pmatrix} y+\varepsilon g(t, y, y'),$$

$$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} y(0)+\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix} y(2\pi/\sigma)=0,$$

where y and g are defined as usual. For this system

$$U(t)=\begin{pmatrix}\sin\sigma t & \cos\sigma t\\ \sigma\cos\sigma t & -\sigma\sin\sigma t\end{pmatrix}$$

and we may take

$$V(t)\begin{pmatrix}\sigma\sin(\sigma t+\theta_0) & \cos(\sigma t+\theta_0)\\ \sigma\cos(\sigma t+\theta_0) & -\sin(\sigma t+\theta_0)\end{pmatrix}$$

where θ_0 will be defined later. The solution to the auxiliary equation has the form $x(t)=c_1\sigma^{-1}\sin(\sigma t)+c_2\sigma^{-1}\cos(\sigma t)$ or $x(t)=\lambda(\varepsilon)\sigma^{-1}\sin(\sigma t+\theta(\varepsilon))+O(\varepsilon)$ where $\lambda(\varepsilon)=\lambda_0+O(\varepsilon)$ and $\theta(\varepsilon)=\theta_0+O(\varepsilon)$. In this problem we have

$$H_1(\lambda, \theta, \varepsilon)=\int_0^{2\pi/\sigma} \{(1-\lambda^2\sigma^{-2}\sin^2(\sigma t+\theta))\lambda\cos(\sigma t+\theta)+b\sigma\cos(\sigma t+\alpha)$$

$$-a\lambda\sin(\sigma t+\theta)\}\cos(\sigma t+\theta_0)\,dt,$$

$$H_1(\lambda_0, \theta_0, 0)=\frac{-\pi}{4\sigma^3}\{\lambda_0^3-4\lambda_0\sigma^2-4b\sigma^3\cos(\alpha-\theta_0)\},$$

$$H_2(\lambda, \theta, \varepsilon)=\int_0^{2\pi/\sigma} \{\cdots\}\sin(\sigma t+\theta_0)\,dt,$$

$$H_2(\lambda_0, \theta_0, 0)=\frac{-\pi}{\sigma}\{a\lambda_0+b\sigma\sin(\alpha-\theta_0)\},$$

$$\frac{\partial H(\lambda_0, \theta_0, 0)}{\partial(\lambda, \theta)}=\begin{pmatrix}\dfrac{-\pi}{4\sigma^3}(3\lambda_0^2+4\sigma^2) & \dfrac{-\pi}{4\sigma^3}(-4b\sigma^2\sin(\alpha-\theta_0))\\ \dfrac{-\pi}{\sigma}(a) & \dfrac{-\pi}{\sigma}(-b\sigma\cos(\alpha-\theta_0))\end{pmatrix}.$$

Now $H(\lambda_0, \theta_0, 0)=0$ has a solution when $a \ll b$ and $b \ll 1$ of the form $\lambda_0^2 \simeq 4\sigma^2$ and $\theta_0 \simeq \alpha$. For these values of λ_0 and θ_0 the Jacobian determinant becomes a nonzero term times

$$3\lambda_0^2+4\sigma^2-4a\sigma \tan(\alpha-\theta_0) \simeq 16\sigma^2.$$

Hence for σ different from zero, $a \ll b$ and $b \ll 1$, then it follows from Theorem 5.3 that equation (6.2) has a $2\pi/\sigma$-periodic solution of the form $x(t)=\lambda(\varepsilon)\sigma^{-1}\sin(\sigma t+\theta(\varepsilon))+O(\varepsilon)$ where $\lambda(\varepsilon)\simeq 2\sigma+O(\varepsilon)$ and $\theta(\varepsilon)\simeq\alpha+O(\varepsilon)$.

Example 3.

$$(6.3)\qquad \begin{gathered} y'=\begin{pmatrix}0&1&0\\0&0&1\\0&0&1\end{pmatrix} y+\varepsilon\begin{pmatrix}f_1(t,y,y')\\f_2(t,y,y')\\f_3(t,y,y')\end{pmatrix}, \\ \begin{pmatrix}1&-1&0\\0&0&0\\1&0&-1\end{pmatrix} y(0)+\begin{pmatrix}0&0&0\\1&-1&0\\0&0&0\end{pmatrix} y(1)=0. \end{gathered}$$

Here, $p=1$ and $q=1$. For this system we have $U(t)=\operatorname{col}(e^t, e^t, e^t)$ and $v(t)=(-1, t, 1-t)$. The bifurcation equation has the form

$$H(\alpha,\varepsilon)=\int_0^1 \{-f_1(t,x^*,x^{*\prime})+tf_2(t,x^*,x^{*\prime}) \\ +(1-t)f_3(t,x^*,x^{*\prime})\}\,dt.$$

If it happens that $f_1=f_2=f_3$, then we have $H(\alpha,\varepsilon)\equiv 0$, and in this situation, for each ε small, we have a one parameter family of the form $x(t)=\operatorname{col}(ce^t, ce^t, ce^t)+O(\varepsilon)$ where $|c|\leqq\rho$ and this family is a solution to equation (6.3).

Example 4.

$$(6.4)\qquad \begin{gathered} x''+x=\varepsilon f(t,x,x',x''), \\ x'(0)=x'(2\pi). \end{gathered}$$

This boundary value problem can be written as the system

$$y'=\begin{pmatrix}0&1\\-1&0\end{pmatrix} y+\varepsilon g(t,y,y'),$$

$$(0,1)y(0)+(0,-1)y(2\pi)=0$$

where y and g are defined as usual. For this system

$$U(t)=\begin{pmatrix}\sin t&\cos t\\\cos t&-\sin t\end{pmatrix}$$

and $V(t)=(\sin t, \cos t)$. This example is nonselfadjoint and in fact p and q are not equal, i.e. $p=2$ and $q=1$. The solution to the auxiliary equation is of the form. $x(t)=\alpha_1 \sin t+\alpha_2\cos t+O(\varepsilon)$. Here

$$H(\alpha,\varepsilon)=\int_0^{2\pi} \cos t f(t,x,x',x'')\,dt.$$

If $f=g\cos t$ with $g(t,x,x',x'')\geqq 0$ and not identically zero, then $H(\alpha,\varepsilon)\neq 0$ and there would be no solutions to the boundary value problem (6.4). However, if $f(t, x, x', x'')=x^3+\sin t$, then

$$H(\alpha, 0)=(3\pi/4)(\alpha_1^2\alpha_2+\alpha_2^3),$$

$$\partial H(\alpha, 0)/\partial\alpha=((3\pi/2)\alpha_1\alpha_2 \quad (3\pi/4)(\alpha_1^2+3\alpha_2^2)).$$

If $\alpha_1\neq 0$ and $\alpha_2=0$, then $H(\alpha, 0)=0$ and $\partial H(\alpha, 0)/\partial\alpha$ has maximal rank 1. Hence, by Theorem 5.3 the boundary value problem (6.4) has a solution of the form $x(t)=\alpha_1 \sin t+O(\varepsilon)$ where $\alpha_1\neq 0$.

Acknowledgments. The author wishes to acknowledge the many helpful conversations held with Professor L. Cesari regarding the subject of this paper and the many useful suggestions made by the referees.

REFERENCES

[1] L. CESARI, *Functional analysis and periodic solutions of nonlinear differential equations*, Contributions to Differential Equations 1, John Wiley, New York, 1963, pp. 149–187.

[2] E. A. CODDINGTON AND N. LEVINSON, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[3] R. H. COLE, *Theory of Ordinary Differential Equations*, Appleton-Century-Crofts, New York, 1968.

[4] R. A. GAMBILL AND J. K. HALE, *Subharmonic and ultraharmonic solutions for weakly nonlinear solutions*, J. Rational Mech. Anal., 5 (1956), pp. 353–394.

[5] J. K. HALE, *Oscillations in Nonlinear Systems*, McGraw-Hill, New York, 1963.

[6] ———, *Ordinary Differential Equations*, Wiley–Interscience, New York, 1969.

[7] J. MAWHIN, *Degré topologique et solutions périodique des systèmes différentiels non linéaires*, Bull. Roy. Soc. Liège, 38 (1969), pp. 308–398.

[8] R. K. NAGLE, *Boundary value problems for nonlinear ordinary differential equations*, Thesis, Univ. of Michigan, Ann Arbor, June 1975.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0015 $01.00/0

EXISTENCE AND APPROXIMATION OF WEAK SOLUTIONS OF NONLINEAR DIRICHLET PROBLEMS WITH DISCONTINUOUS COEFFICIENTS*

JOSEPH W. JEROME†

Abstract. The Dirichlet problems discussed in this paper arise when implicit time approximation methods are employed in perturbed two-phase Stefan problems. The discontinuity in the enthalpy h across the free boundary interface of the two phases appears in the Dirichlet problems as a term $h(U)$, where h is discontinuous at 0. Two major results are presented, viz., an existence theorem, making use of pseudomonotone operators, and an approximation theorem, utilizing solutions U_ε of appropriately smoothed Dirichlet problems corresponding to smoothings h_ε of h. In the special case of homogeneous boundary conditions, an alternative approach, making use of results of Brezis–Strauss together with "a priori" estimates and Leray–Schauder degree theory, gives existence of solutions.

1. Introduction. In this paper we demonstrate the existence and approximation of weak solutions U of nonlinear Dirichlet problems, on bounded domains $\Omega \subset R^N$, of the form

$$\text{(1.1)} \qquad \begin{aligned} &\text{(i)} \quad LU + h(U) + g(U) \ni f, \\ &\text{(ii)} \quad U - W \in H_0^1(\Omega), \qquad W \in H^1(\Omega), \end{aligned}$$

where L is a (formally) self-adjoint elliptic operator of second order determining a strongly coercive quadratic form on $H_0^1(\Omega)$, h is a monotone increasing function, *discontinuous* at 0, g is a certain Lipschitz function, not assumed convex or monotone, and $f \in H^{-1}(\Omega)$. The precise hypotheses are presented in § 2. We note here that g is assumed to admit the decomposition $g = g_1 + g_2$, where $g_1(\lambda)\lambda \geqq 0$ and $\|g_2\|_{\text{Lip}}$ is strictly less than the smallest eigenvalue of L.

Our interest in such problems arose directly from attempts at constructing weak solutions of free boundary diffusion equations which possess discontinuous diffusion coefficients. Specifically, if we consider the two-phase Stefan problem,

$$\text{(1.2)} \qquad \frac{\partial u}{\partial t} - \nabla \cdot (k(u)\nabla u) + a(u) = f,$$

on a space-time domain $\Omega \times (0, T_0)$ with prescribed initial and time invariant boundary conditions, and prescribed enthalpy discontinuity across the free boundary, where the diffusion coefficient k is a positive function with compact range closure in $(0, \infty)$ which is discontinuous at 0 and a is a certain body heating Lipschitz function, then the Kirchhoff transformation,

$$U = K(u) = \int_0^u k(\lambda)\, d\lambda,$$

transforms (1.2) into the form,

$$\text{(1.3)} \qquad \frac{\partial h(U)}{\partial t} - \Delta U + g(U) = f,$$

* Received by the editors April 30, 1976, and in revised form November 15, 1976.

† Department of Mathematics and the Technological Institute, Northwestern University, Evanston, Illinois 60201. This research was supported at Oxford University by a grant from the British Science Research Council and by National Science Foundation under Grant MPS74-02292 A01.

where $h'(\lambda)=1/(k(K^{-1}(\lambda)))$, $\lambda\neq 0$, and h has a prescribed (enthalpy) discontinuity at 0; g satisfies

$$g(\lambda)=a(K^{-1}(\lambda)).$$

If an implicit time discretization is employed in (1.3), one obtains the (finite) sequence,

$$[h(U_{n+1})-h(U_n)]/\Delta t-\Delta U_{n+1}+g(U_{n+1})=f, \tag{1.4}$$

of nonlinear elliptic boundary value problems where $U_0=K(u_0)$ is specified. Equation (1.4) is of course a special case of (1.1) where $L=-\Delta$. It will be shown in the paper, [9], and has already been announced in the survey paper [7], how a weak solution of (1.3), in the sense of Oleinik [12], can be approximated by $H^1(\Omega)$-valued piecewise linear or step functions constructed from the solutions of (1.4).

Our major results are Theorem 2.2 which states that solutions of (1.1) exist and Theorem 2.4 which states that such solutions can be approximated by solutions of certain smoothed problems. In the applications to nonlinear equations of evolution, we see that (1.4) has a unique solution for an arbitrary Lipschitz continuous function g, provided $\Delta t\leqq[\|g\|_{\mathrm{Lip}}\sup k]^{-1}$ Theorem 2.4 is essential to the derivation of stability relations satisfied by the solutions of (1.4) (cf. [9, Thm. 3.1]).

Our methods of proof make use of the theory of pseudomonotone operators, as developed by Brezis [4], for the proof of Theorem 2.2 and "a priori" estimates, derived by application of the Brouwer fixed point theorem to eigenspaces of L, for the proof of Theorem 2.4. Such estimates permit, via Galerkin approximation of the smoothed solution, the determination of a fixed sphere in $H_0^1(\Omega)$, which contains a solution of the smoothed problem for each value of the smoothing parameter. Any weak limit point V of this family satisfies the property that $U=V+W$ is a solution of (1.1); the chief technical difficulty here is the definition of $h(U)$ on subsets of Ω on which U vanishes.

If g is a convex, monotone increasing function then, at least in the case of the smoothed problem it is known that the growth restriction, that g be Lipschitz on R, can be relaxed. The methods here make essential use of the maximum principle. See, e.g., Parter [13], Keller [10] and Schryer [16]. These approaches would not appear to be successful, however, in the general unsmoothed case.

Another standard approach to nonlinear elliptic boundary value problems is the method of linearization, or effective linearization, combined with duality. An example of this approach is furnished by Rosenzweig [14] who considers the Sobolev Banach spaces, thereby handling polynomial growth for g. This requires, however, coefficient regularity not satisfied in the present problem. It is possible, though, that this approach, when combined with smoothing, could prove successful in certain cases.

We note that the results derived in this paper are valid for an arbitrary bounded open set Ω in R^N; this general approach has been motivated by the book of Lions [11] (cf. also Dubinski [5]).

2. The major results. Let Ω be a bounded region in R^N, $N\geqq 1$. The Sobolev Hilbert spaces $H^1(\Omega)$ and $H_0^1(\Omega)$ will have their usual meaning as the completion of $C^\infty(\Omega)$ and $C_0^\infty(\Omega)$, respectively, in the norm determined by the inner product

$$(u,v)_{H^1}=\int_\Omega uv+\int_\Omega \nabla u\cdot\nabla v. \tag{2.1}$$

All functions are real-valued and $H^{-1}(\Omega)$ is the (topological) dual of $H_0^1(\Omega)$.

We shall consider symmetric linear elliptic operators L of second order of the form,

$$L = \sum_{\substack{0 \leqq |\alpha| \leqq 1 \\ 0 \leqq |\beta| \leqq 1}} (-1)^{|\alpha|} D^\alpha (a_{\alpha\beta} D^\beta), \tag{2.2}$$

where $a_{\alpha\beta} \in L^\infty(\Omega)$, $0 \leqq |\alpha|, |\beta| \leqq 1$, L is understood in the sense of distributions[1] and where it is assumed that the bilinear form $B(\cdot, \cdot)$ on $H^1(\Omega)$,

$$B(u, v) = \int_\Omega \sum_{\substack{0 \leqq |\alpha| \leqq 1 \\ 0 \leqq |\beta| \leqq 1}} a_{\alpha\beta} D^\alpha u D^\beta v, \tag{2.3}$$

is a symmetric bilinear form which determines a norm on $H_0^1(\Omega)$ equivalent to that of the Sobolev norm:

$$C(u, u)_{H^1} \leqq B(u, u), \quad \text{for all } u \in H_0^1(\Omega), \quad \text{for some } C > 0. \tag{2.4}$$

From (2.4) and the fact that $B(\cdot, \cdot)$ is a continuous bilinear form on $H_0^1(\Omega)$ we conclude, by a standard application of the Lax–Milgram lemma [2, p. 30], that L is a continuous bijection of $H_0^1(\Omega)$ onto its dual $H^{-1}(\Omega)$ given by, for each fixed $v \in H_0^1(\Omega)$,

$$\langle Lv, u \rangle = B(u, v), \quad \text{for all } u \in H_0^1(\Omega). \tag{2.5}$$

Denote by $\tilde{L}$ the restriction of L to $D_{\tilde{L}} = L^{-1}(L^2(\Omega))$. Then, for all $v \in D_{\tilde{L}}$,

$$(\tilde{L}v, u)_{L^2} = B(u, v), \quad \text{for all } u \in H_0^1(\Omega). \tag{2.6}$$

Equation (2.6) follows by another application of the Lax–Milgram theorem to the continuous linear functional $(f, \cdot)_{L^2}$ on $H_0^1(\Omega)$, where $f \in L^2(\Omega)$. The identity,

$$B(\tilde{L}^{-1}f, \tilde{L}^{-1}f) = (f, \tilde{L}^{-1}f)_{L^2},$$

together with (2.4) and the Schwarz inequality, show that $\tilde{L}^{-1}$ is a continuous mapping of $L^2(\Omega)$ into $H_0^1(\Omega)$. Since the injection $H_0^1(\Omega) \to L^2(\Omega)$ is compact [1, p. 99], it follows that $\tilde{L}^{-1}$ is compact when viewed as an operator from $L^2(\Omega)$ into itself. We summarize this as

LEMMA 2.1. *The restriction of L^{-1} to $L^2(\Omega)$ is a compact linear operator on $L^2(\Omega)$.*

It follows from (2.4), Lemma 2.1 and the spectral theory for symmetric compact operators that $\tilde{L}^{-1}$, and hence $\tilde{L}$, has a complete orthonormal sequence of eigenfunctions in $L^2(\Omega)$ which are orthogonal and complete in $H_0^1(\Omega)$. This standard fact is documented, for example, in [8, Thm. 3.2].

Now let h be a strictly monotone function, discontinuous at 0, defined by,

$$\begin{aligned} &\text{(i)} \quad h'(\lambda) = \theta(\lambda), \qquad \lambda \neq 0, \\ &\text{(ii)} \quad h(0+) - h(0-) = b > 0, \\ &\text{(iii)} \quad h(0-) = 0, \end{aligned} \tag{2.7}$$

where θ is a positive function with compact range closure in $(0, \infty)$ which is continuous on $R - \{0\}$ with positive right and left hand limits at 0. h defines a bounded operator H from $L^2(\Omega)$ into $L^2(\Omega)$ by

$$Hf(x) = h(f(x)), \qquad x \in \Omega.$$

Here it is understood, and this is important for the sequel, that $h(f(x))$ is any value in

[1] L is defined precisely by (2.5)

$[0, b]$ such that $h \cdot f$ is measurable if $f(x)=0$. With this convention, Hf is measurable and, in particular, square integrable if f is. Also, H is a strictly monotone operator on $L^2(\Omega)$:

$$(H(f)-H(g), f-g)_{L^2} \geqq \theta_1(f-g, f-g)_{L^2}, \tag{2.8}$$

where $0<\theta_1=\inf\{\theta(\lambda):\lambda \in R-\{0\}\}$. Finally, H defines a mapping from $H^1(\Omega)$ into its dual upon identifying $L^2(\Omega)$ with its dual. In this case,

$$H^1(\Omega) \subset L^2(\Omega) \subset (H^1(\Omega))', \tag{2.9}$$

where the injections are dense and continuous, and

$$\langle H(u), v\rangle = (H(u), v)_{L^2}.$$

Now let g be a Lipschitz function on R of the form,

$$g = g_1 + g_2,$$

where g_1 satisfies the condition,

$$g_1(\lambda)\lambda \geqq 0, \qquad \lambda \in R, \tag{2.10}$$

and where g_2 satisfies a Lipschitz condition,

$$|g_2(x)-g_2(y)| \leqq C_2|x-y|, \tag{2.11}$$

$$0 \leqq C_2 < C + \theta_1 \tag{2.12}$$

where C is the constant of (2.4). The hypothesis (2.12) can be replaced by the stronger hypothesis,

$$0 \leqq C_2 < \omega + \theta_1 \tag{2.13}$$

where ω is the smallest eigenvalue of $\tilde{L}$ since $\omega \leqq C$. The mapping G given by

$$Gf(x) = g(f(x)), \qquad x \in \Omega,$$

defines a continuous mapping on $L^2(\Omega)$.

DEFINITION 2.1. By a *weak solution of the nonlinear Dirichlet problem for the operator* $L+H+G$ *for prescribed* $W \in H^1(\Omega)$ *and* $F \in H^{-1}(\Omega)$ is meant a function $U = V + W$, $V \in H_0^1(\Omega)$, such that,

$$B(V, v) + (H(U), v)_{L^2} + (G(U), v)_{L^2} = F(v) - B(W, v) \tag{2.14}$$

for all $v \in H_0^1(\Omega)$. In particular, $TV \ni F - B(W, \cdot) = F_0$, where $Tv = Lv + H(v+W) + G(v+W)$ maps $H_0^1(\Omega)$ into $2^{H^{-1}(\Omega)}$.

THEOREM 2.2. *Under the stated hypotheses on* Ω, L, h, g *and* W, *there is a solution* U *of* (2.14). U *is unique if* $g_1 \equiv 0$.

Our next result deals with the approximation of U by solutions of smoothed problems. It is necessary in the derivation of stability inequalities (cf. [9, Thm. 3.1]) in the analysis of nonlinear equations of evolution. We shall smooth the enthalpy function h in a precise way to obtain a net j_ε of continuously differentiable Lipschitz functions satisfying $j_\varepsilon(0)=0$, $j_\varepsilon' \geqq \theta_1$ and $j_\varepsilon \to h$ on $R-\{0\}$. Specifically, let

$$\delta_* = |\theta(0+) - \theta(0-)|, \qquad \theta_0 = \theta(0-) \tag{2.15}$$

and define $1 + \delta_* = \delta > 0$. Select $0 < \varepsilon_0 \leqq 1$ such that

$$|\theta(\varepsilon) - \theta(0+)| \leqq \delta/2 \quad \text{if } 0 < \varepsilon < \varepsilon_0.$$

If

$$\varepsilon_1 = \min\left[\frac{2b}{6\theta_0+3\delta}, \frac{8b}{3\delta}\right], \tag{2.16}$$

set $\varepsilon_* = \min(\varepsilon_0, \varepsilon_1)$. Henceforth, we restrict the smoothing parameter ε to the interval $0<\varepsilon \leqq \varepsilon_*$.

Now let ω be defined by, for fixed ε,

$$\begin{aligned} &\text{(i)} \quad \omega(\lambda) = \theta(\lambda), \qquad \lambda < 0, \quad \lambda \geqq \varepsilon, \\ &\text{(ii)} \quad \omega(\lambda) = q(\lambda), \qquad 0 \leqq \lambda \leqq \varepsilon. \end{aligned} \tag{2.17}$$

Here q is the uniquely determined quadratic polynomial on $[0, \varepsilon]$ defined by

$$q(0) = \theta_0, \quad q(\varepsilon) = \theta(\varepsilon), \quad \int_0^\varepsilon q(\lambda)\, d\lambda = b. \tag{2.18}$$

A routine calculation shows that q is given explicitly by,

$$q(\lambda) = q_0 + q_1\lambda + q_2\lambda^2, \tag{2.19}$$

where

$$\begin{aligned} &q_0 = \theta_0, \qquad q_1 = 2[3b - 2\varepsilon\theta_0 - \varepsilon\theta(\varepsilon)]/\varepsilon^2, \\ &q_2 = \{(\theta(\varepsilon) - \theta_0)\varepsilon - 2[3b - 2\varepsilon\theta_0 - \varepsilon\theta(\varepsilon)]\}/\varepsilon^3. \end{aligned} \tag{2.20}$$

We now define $j = j_\varepsilon$ by,

$$j(\lambda) = \int_0^\lambda \omega(t)\, dt. \tag{2.21}$$

LEMMA 2.3. *ω is a continuous, positive Lipschitz function and j is a continuously differentiable Lipschitz function on R satisfying*

$$j(\lambda)\lambda \geqq 0, \qquad \lambda \in R. \tag{2.22}$$

Moreover, the nets $\{\omega = \omega_\varepsilon\}$ and $\{j = j_\varepsilon\}$ converge uniformly on compact subsets of $R - \{0\}$ to θ and h, respectively. Finally, q satisfies the inequality

$$q(\lambda) \geqq \inf(\theta_0, \theta(\varepsilon)), \qquad 0 < \lambda \leqq \varepsilon. \tag{2.23}$$

In particular, $j'(\lambda) \geqq \theta_1$, $\lambda \in R$.

Proof. ω is clearly a continuous Lipschitz function; the positivity of ω will follow from the inequality (2.23). We shall show that q is concave on $[0, \varepsilon]$, i.e., $q_2 \leqq 0$. Using the inequalities,

$$\varepsilon \leqq \varepsilon_* \leqq \frac{2b}{6\theta_0 + 3\delta}, \qquad \theta(\varepsilon) \leqq \theta_0 + 3\delta/2,$$

we conclude that $\varepsilon q_1 \geqq 4b/\varepsilon$. Thus,

$$\begin{aligned} \varepsilon^2 q_2 &= [\theta(\varepsilon) - \theta_0] - \varepsilon q_1 \\ &\leqq \left(\frac{3\delta}{2}\right) - \frac{4b}{\varepsilon} \end{aligned}$$

which, together with the inequality,

$$\varepsilon \leqq \varepsilon_* \leqq \frac{8b}{3\delta},$$

yields $q_2 \leqq 0$. Thus, q is concave and its minimum is achieved at one of the endpoints of $[0, \varepsilon]$, proving (2.23). The statements concerning j, including (2.22), are now clear as is the uniform convergence of ω_ε to θ on compact subsets of $R-\{0\}$. To prove the uniform convergence of $\{j=j_\varepsilon\}$, let $A \subset R-\{0\}$, A compact. Since j agrees with h on $(-\infty, 0)$, we may assume $A \subset R^+$. Choose ε so that $(0, \varepsilon) \cap A = \emptyset$. Then, for $t \in A$, we have

$$h(t)-j(t)=\left[b+\int_0^t \theta(\lambda)\,d\lambda\right]-\left[\int_0^\varepsilon q(\lambda)\,d\lambda+\int_\varepsilon^t \theta(\lambda)\,d\lambda\right]$$
$$=\int_0^\varepsilon \theta(\lambda)\,d\lambda$$

which can be made arbitrarily small if ε is sufficiently small. This completes the proof of the lemma.

Note that a mapping $J=J_\varepsilon$, continuous on $L^2(\Omega)$ and $H^1(\Omega)$, is determined in the usual way by j_ε. Consider the mapping $T_\varepsilon: H_0^1(\Omega) \to H^{-1}(\Omega)$ given by,

$$T_\varepsilon v = Lv + J_\varepsilon(v+W) + G(v+W), \qquad J_\varepsilon u = j_\varepsilon(u). \tag{2.24}$$

THEOREM 2.4. *Given $F_0 \in H^{-1}(\Omega)$, there is a sphere $\mathcal{B} \subset H_0^1(\Omega)$, with radius independent of $0<\varepsilon \leqq \varepsilon_*$, such that $\mathcal{B}$ contains at least one solution V_ε of the equation $T_\varepsilon V_\varepsilon = F_0$. For any sequence $\varepsilon_\nu \to 0$, there is a subsequence ε_{ν_j} such that $V_{\varepsilon_{\nu_j}} \to V$ in $L^2(\Omega)$ and $J_{\varepsilon_{\nu_j}}(V_{\varepsilon_{\nu_j}}+W) \rightharpoonup H(V+W)$ in $L^2(\Omega)$.*

3. Existence and a priori estimates. We begin this section by furnishing a proof for Theorem 2.2.

Proof of Theorem 2.2. The mapping $V \to AV = H(V+W)$ is a (multi-valued) maximal monotone operator from $H_0^1(\Omega)$ into $2^{L^2(\Omega)}$ (and thus into $2^{H^{-1}(\Omega)}$). Indeed, A is the subdifferential $A=\partial\varphi$ of the finite-valued continuous, convex functional,

$$\varphi(V)=\int_\Omega l(V(x)+W(x))\,dx;$$

here l is the (convex) primitive of h satisfying $l(0)=0$. But the subdifferential of any lower semicontinuous proper convex functional is maximal monotone, so that A is maximal monotone.

We claim that the operator

$$V \to BV = LV + G(V+W)$$

is *pseudomonotone* from $H_0^1(\Omega)$ into $H^{-1}(\Omega)$, i.e., B satisfies,

(3.1i) B is bounded;

(3.1ii)
$$V_i \rightharpoonup V \text{ in } H_0^1(\Omega) \text{ and } \limsup_{i\to\infty} \langle BV_i, V_i - V\rangle \leqq 0$$
$$\Rightarrow \liminf_{i\to\infty} \langle BV_i, V_i - Z\rangle \geqq \langle BV, V-Z\rangle \ \forall Z \in H_0^1(\Omega).$$

Indeed, L is a continuous linear mapping of $H_0^1(\Omega)$ into $H^{-1}(\Omega)$ and G is a bounded mapping of $H_0^1(\Omega)$ into $L^2(\Omega)$ and hence into $H^{-1}(\Omega)$ so that (3.1i) holds. Because of the compact injection of $H_0^1(\Omega)$ into $L^2(\Omega)$, we have even the stronger implication,

$$\text{(3.1iii)}\qquad \begin{aligned} &V_i \rightharpoonup V \text{ in } H_0^1(\Omega) \text{ and } \limsup_{i\to\infty} \langle BV_i, V_i - V\rangle \leqq 0 \\ &\Rightarrow V_i \to V \text{ in } H_0^1(\Omega), \end{aligned}$$

which clearly implies (3.1ii). To verify (3.1iii), note that

$$\limsup_{i\to\infty} \langle BV_i, V_i - V\rangle \leqq 0 \Rightarrow \limsup_{i\to\infty} \langle LV_i, V_i - V\rangle \leqq 0$$

since $G(V_i + W) \to G(V + W)$ in $L^2(\Omega)$; moreover, it follows from (2.4) and (2.5) that

$$\liminf_{i\to\infty} \langle LV_i - LV, V_i - V\rangle \geqq 0,$$

and hence that

$$\liminf_{i\to\infty} \langle LV_i, V_i - V\rangle \geqq 0.$$

Altogether, then, $\lim_{i\to\infty} \langle LV_i, V_i - V\rangle = 0$, so that $\lim_{i\to\infty} \langle LV_i - LV, V_i - V\rangle = 0$, i.e., $V_i \to V$ in $H_0^1(\Omega)$. Thus, B is pseudomonotone. Now $A+B$ is coercive by (2.10)–(2.12):

$$\langle AV + BV, V\rangle \geqq c\|V\|^2_{H^1(\Omega)} - [\bar{C} + b(\text{meas}\,\Omega)^{1/2}]\|V\|_{L^2(\Omega)},$$

where c and $\bar{C}$ are given by (3.2) and (3.10) below. We omit the details because of the similarity to Proposition 3.1. The surjectivity of $A+B$ now follows from [4, Théorème 1], where it is proved that the coercive sum of a maximal monotone and pseudomonotone operator is surjective.

If $g_1 \equiv 0$ and U_1 and U_2 are solutions of (2.14) then, with $U_1 - U_2 = U \in H_0^1(\Omega)$,

$$\begin{aligned} 0 &= (H(U_1) - H(U_2), U)_{L^2(\Omega)} + B(U, U) + (G(U_1) - G(U_2), U)_{L^2(\Omega)} \\ &\geqq (C - C_2 + \theta_1)\|U\|^2_{L^2(\Omega)} \geqq 0. \end{aligned}$$

In particular, $U = 0$ and uniqueness holds. This completes the proof of Theorem 2.1.

We begin the analysis of "a priori" estimates with a uniform coerciveness estimate.

PROPOSITION 3.1. *Let T_ε be defined by* (2.24). *Then there exists a constant $\bar{C}$, independent of ε, such that*

$$\text{(3.2)}\qquad \frac{\langle T_\varepsilon(v), v\rangle}{\|v\|_{H^1}} \geqq c\|v\|_{H^1} - \bar{C} \quad \textit{for all } v \in H_0^1(\Omega),$$

where c is given explicitly by

$$c = \min(C, C + \theta_1 - C_2).$$

LEMMA 3.2. *There exists a positive constant C_0* such that

$$\text{(3.3)}\qquad 0 < q(t) \leqq C_0/\varepsilon, \quad \textit{for } 0 \leqq t \leqq \varepsilon,$$

and, for $v \in H_0^1(\Omega)$, we have

$$\text{(3.4)}\qquad |(j(v + W) - j(v), v)_{L^2}| \leqq C_1(\|W\|_{L^2} + [\text{meas}\,\Omega]^{1/2})\|v\|_{L^2}$$

where

$$\text{(3.5)}\qquad C_1 = \max(C_0, \sup(\theta(t): t \in R - \{0\})); \quad \textit{here } j = j_\varepsilon.$$

Proof of Lemma 3.2. Condition (3.3) follows immediately from (2.20). The verification of (3.4) involves the estimation of $|j(v(x)+W(x))-j(v(x))|$ for x fixed in Ω. A careful distinction of cases with respect to the location of $v(x)+W(x)$ and $v(x)$, relative to the interval $[0, \varepsilon]$, yields the estimate

$$|j(v(x)+W(x))-j(v(x))| \leqq C_1(1+|W(x)|). \tag{3.6}$$

Here we have used the fact that, for any $0 \leqq t_2 < t_1 \leqq \varepsilon$, the estimate

$$\begin{aligned}|j(t_1)-j(t_2)| &= \left|\frac{j(t_1)-j(t_2)}{t_1-t_2}\right|(t_1-t_2) \\ &\leqq (C_0/\varepsilon)\varepsilon = C_0\end{aligned}$$

holds. Condition (3.4) follows immediately from (3.6).

Proof of Proposition 3.1. To verify (3.2), note that

$$\begin{aligned}\langle T_\varepsilon(v), v\rangle = [B(v, v) + (j(v), v)_{L^2} + (g_2(v), v)_{L^2}] + (g_1(v), v)_{L^2} \\ +(j(v+W)-j(v), v)_{L^2} + (g(v+W)-g(v), v)_{L^2}.\end{aligned} \tag{3.7}$$

Now the last term is estimated similarly to (3.4):

$$|(g(v+W)-g(v), v)_{L^2}| \leqq C_2\|W\|_{L^2}\|v\|_{L^2}, \tag{3.8}$$

if C_2 is a Lipschitz constant for g. Noting the obvious consequence,

$$(g_1(v), v)_{L^2} \geqq 0, \tag{3.9}$$

of (2.10) we have, altogether, from (2.12), (3.4), (3.8) and (3.9), the inequality (3.2) where $\bar{C}$ is given by

$$\bar{C} = (C_2+C_1)\|W\|_{L^2} + [\text{meas } \Omega]^{1/2}(C_1+|g_2(0)|). \tag{3.10}$$

This concludes the proof of Proposition 3.1.

Let S be any finite dimensional subspace of $H_0^1(\Omega)$ and let P be the projection onto S which maps an element v in $H_0^1(\Omega)$ onto the unique element $Pv = u$ in S, which is closest to v in the norm $[B(\cdot, \cdot)]^{1/2}$:

$$[B(v-Pv, v-Pv)]^{1/2} = \inf_{s \in S}[B(v-s, v-s)]^{1/2}. \tag{3.11}$$

Pv is uniquely determined since the norm $[B(\cdot, \cdot)]^{1/2}$ is strictly convex; in fact, with the inner product $B(\cdot, \cdot)$ the set $H_0^1(\Omega)$ is a Hilbert space which uniquely defines projections onto closed subspaces. Now we define the continuous mapping $P^t: H^{-1}(\Omega) \to H^{-1}(\Omega)$ by

$$\langle P^t F, v\rangle = \langle F, Pv\rangle, \qquad v \in H_0^1(\Omega). \tag{3.12}$$

If $\{S_n\}$ is the sequence of finite dimensional subspaces of $H_0^1(\Omega)$ spanned by the first n eigenfunctions of $\tilde{L}$, we denote the corresponding induced projections by $\{P_n\}$.

The coerciveness inequality of Proposition 3.1 plays a fundamental role in the following. S' denotes the dual of S in the sequel.

PROPOSITION 3.3. *Let S be a finite dimensional subspace of $H_0^1(\Omega)$ and let P be defined by* (3.11) *and P^t by* (3.12). *Then the mapping $P^t T_\varepsilon P$ has the property that, for a fixed $F_0 \in H^{-1}(\Omega)$, there is a ball of radius $r = r(T, W, F_0)$ in $H_0^1(\Omega)$ containing a* *solution s of the equation*

$$P^tT_\varepsilon Ps = P^tF_0 \tag{3.13}$$

with the property that r does not depend upon ε nor upon S.[2]

We shall have need of the following lemma [11, p. 53, Lemma 4.3].

LEMMA 3.4. *Let* Γ *be a continuous mapping of real Euclidean space* R^m *into itself such that, for some* $\rho > 0$, *one has*

$$(\Gamma\xi, \xi)_{R^m} \geqq 0, \quad \text{for all } \xi \text{ such that } |\xi| = \rho, \tag{3.14}$$

where, for $\xi = (\xi_1, \cdots, \xi_m)$, $\eta = (\eta_1, \cdots, \eta_m)$,

$$(\xi, \eta)_{R^m} = \sum_{i=1}^m \xi_i\eta_i, \qquad |\xi| = (\xi, \xi)^{1/2}. \tag{3.15}$$

There then exists $\xi \in R^m$, $|\xi| \leqq \rho$, *such that* $\Gamma\xi = 0$.

Proof of Proposition 3.3. Let $F_0 \in H^{-1}(\Omega)$ be fixed and define the mapping $\Lambda: S \to S'$ by

$$\Lambda v = P^t\, T_\varepsilon v - P^tF_0, \qquad v \in S. \tag{3.16}$$

The assertion of Proposition 3.3 requires that there exist $v \in S$ for which $\Lambda v = 0$. Now if the dimension of S is m, then Λ induces a mapping Γ of R^m into R^m as follows. Let $s_1, \cdots, s_m$ denote an orthonormal basis for S and let $\xi \in R^m$. Then $s = \sum_{i=1}^m \xi_is_i \in S$; if $(\Lambda s)(s_i) = \eta_i$, $i = 1, \cdots, m$, define $\Gamma\xi = \eta$.

Now,

$$(\Gamma\xi, \xi)_{R^m} = \langle\Lambda(s), s\rangle = \langle T_\varepsilon(s), s\rangle - \langle F_0, s\rangle, \tag{3.17}$$

and by the orthonormality of the basis, we have

$$|\xi| = \|s\|_{H^1}, \qquad s = \sum_{i=1}^m \xi_is_i. \tag{3.18}$$

Thus, we obtain from (3.17), (3.18) and (3.2),

$$(\Gamma\xi, \xi)_{R^m} \geqq \|s\|_{H^1}[c\|s\|_{H^1} - \bar{C} - \|F_0\|_{H^{-1}}],$$

which is clearly nonnegative if

$$|\xi| = \|s\|_{H^1} = \rho = [\bar{C} + \|F_0\|_{H^{-1}}]/c. \tag{3.19}$$

Now the continuity of T_ε implies that of Λ, which in turn implies the continuity of Γ. It follows from Lemma 3.4 that there exists $\xi \in R^m$, $|\xi| \leqq \rho$, such that $\Gamma\xi = 0$. Here ρ is given by (3.19). By construction, $\Lambda s = 0$ where s is given by (3.18). s is a solution of (3.13) and $\|s\|_{H^1} \leqq \rho$. If we set $r(T, W, F_0) = \rho$, Proposition 3.3 follows.

PROPOSITION 3.5. *Given* $F_0 \in H^{-1}(\Omega)$, *there is a solution* V_ε *of the equation*

$$T_\varepsilon V_\varepsilon = F_0, \tag{3.20}$$

satisfying $\|V_\varepsilon\|_{H^1} \leqq \rho$. *Here* ρ *is defined by* (3.19).

Proof. Let $\mathscr{B}$ be the closed ball of radius ρ in $H_0^1(\Omega)$ centered at 0 and let S_n be the sequence of eigenspaces of $\tilde{L}$. Now denote by $\psi \in S$ the solution in $\mathscr{B}$, guaranteed by Proposition 3.3, satisfying

$$P_n^tT_\varepsilon P_n\psi_n = P_n^tF_0. \tag{3.21}$$

[2]It follows from Proposition 3.1 that *every* solution s lies in the ball of radius r.

Since $\mathscr{B}$ is weakly compact and since the injection $H_0^1(\Omega) \to L^2(\Omega)$ is compact, there is a subsequence ψ_{nk} satisfying, for some $V_\varepsilon \in \mathscr{B}$,

$$\text{(3.22)} \qquad \begin{aligned} &\text{(i)} \quad \psi_{n_k} \rightharpoonup V_\varepsilon \quad (\text{in } H_0^1(\Omega)), \\ &\text{(ii)} \quad \psi_{n_k} \to V \quad (\text{in } L^2(\Omega)). \end{aligned}$$

Now L is a continuous linear mapping of $H_0^1(\Omega)$ onto $H^{-1}(\Omega)$ and hence weakly continuous [6, p. 422]. Thus $L\psi_{n_k} \rightharpoonup LV_\varepsilon$. J and G are continuous mappings on $L^2(\Omega)$; thus,

$$(J+G)(\psi_{n_k}) \to (J+G)(V_\varepsilon) \quad (\text{in } L^2(\Omega))$$

and also in $H^{-1}(\Omega)$, under the usual definition,

$$\langle (J+G)(u), v\rangle = ((J+G)(u), v)_{L^2}.$$

In particular, $T_\varepsilon \psi_{n_k} \rightharpoonup T_\varepsilon V_\varepsilon$ in $H^{-1}(\Omega)$. It remains to show that $T_\varepsilon\psi_{n_k} - F_0 \rightharpoonup 0$. Now $\langle T_\varepsilon \psi_{n_k} - F_0, P_{n_k} v\rangle = 0$ for all $v \in H_0^1(\Omega)$. Thus,

$$\text{(3.23)} \qquad \begin{aligned} \langle T_\varepsilon \psi_{n_k} - F_0, v\rangle = {} & B(\psi_{n_k}, v - P_{n_k}v) + (J(\psi_{n_k} + W), v - P_{n_k}v)_{L^2} \\ & + (G(\psi_{n_k} + W), v - P_{n_k}v)_{L^2} - \langle F_0, v - P_{n_k}v\rangle. \end{aligned}$$

Now the term $B(\psi_{n_k}, v - P_{n_k}v)$ is zero by the projection theorem; also

$$\text{(3.24)} \qquad \|v - P_{n_k}v\|_{H^1} \to 0$$

by the completeness of the eigenfunctions. It follows from the identification of $H_0^1(\Omega)$ with its second dual and from (3.23), (3.24) and the boundedness of the mappings J and G that $T_\varepsilon \psi_{nk} \rightharpoonup F_0$. Equation (3.20) follows and the proposition is established.

Remark. We note that Proposition 3.5 establishes the first part of Theorem 2.4. The remainder is established in the following section.

4. Convergence of the smoothing approximations. It is the main purpose of this section to prove the convergence properties described in Theorem 2.4.

Proof of Theorem 2.4. Since the solutions V_ε, $0 < \varepsilon \leqq \varepsilon_*$ of (3.20) form a bounded set in $H_0^1(\Omega)$, there is a sequence V_{ε_ν} satisfying

$$\text{(4.1)} \qquad \begin{aligned} &\text{(i)} \quad V_{\varepsilon_\nu} \rightharpoonup V \quad (\text{in } H_0^1(\Omega)), \\ &\text{(ii)} \quad V_{\varepsilon_\nu} \to V \quad (\text{in } L^2(\Omega)). \end{aligned}$$

We may assume without loss of generality, that V_{ε_ν} converges pointwise a.e. to V by taking a subsequence if necessary. Now for each ε_ν,

$$\text{(4.2)} \qquad J_{\varepsilon_\nu}(V_{\varepsilon_\nu} + W) = F_0 - LV_{\varepsilon_\nu} - G(V_{\varepsilon_\nu} + W)$$

and the right hand side of (4.2) is weakly convergent in $H^{-1}(\Omega)$ by the weak continuity of L and G. The sequence defined by the left hand side of (4.2) is bounded in $L^2(\Omega)$; indeed by (3.6) we have

$$\text{(4.3)} \qquad \begin{aligned} \int_\Omega [J_{\varepsilon_\nu}(V_{\varepsilon_\nu} + W)]^2 &\leqq C_1^2 \int_\Omega [1 + |V_{\varepsilon_\nu} + W|]^2 \\ &\leqq 4C_1^2[\text{meas}\,\Omega + \rho^2 + \|W\|_{L^2}^2]. \end{aligned}$$

Thus, there is a subsequence weakly convergent in $L^2(\Omega)$ (and hence in $H^{-1}(\Omega)$) to an $L^2(\Omega)$ function not depending upon the particular subsequence chosen since the entire sequence $J_{\varepsilon_\nu}(V_{\varepsilon_\nu}+W)$ is weakly convergent in $H^{-1}(\Omega)$ as remarked above. In particular, by choosing subsequences if necessary, we may assume (4.1) together with

$$J_{\varepsilon_\nu}(U_{\varepsilon_\nu}) \rightharpoonup Y \quad (\text{in } L^2(\Omega)); \tag{4.4}$$

here we have written $U_{\varepsilon_\nu}=V_{\varepsilon_\nu}+W$.

Now let $\Omega_0=\{x\in\Omega: U(x)=0\}$, $U=V+W$. We shall show that

$$Y(\Omega_0)\subset[0,b], \tag{4.5}$$

up to sets of measure zero. Indeed, if (4.5) fails to hold, there is a set $\Omega_*\subset\Omega_0$ of positive measure satisfying

$$Y(x)\geqq b+\gamma, \qquad x\in\Omega_* \tag{4.6i}$$

for some $\gamma>0$, or

$$Y(x)\leqq -\gamma, \qquad x\in\Omega_*. \tag{4.6ii}$$

We may assume, using Egoroff's theorem if necessary [15, p. 72], that $U_{\varepsilon_\nu} \xrightarrow{\text{uniformly}} 0$ on Ω_*. Suppose that (4.6i) holds and let Φ be the characteristic function of Ω_*. Then $J_{\varepsilon_\nu}(U_{\varepsilon_\nu})\Phi$ is weakly convergent in $L^2(\Omega)$ to $Y\Phi$ and, by the lower semicontinuity of the norm with respect to weak convergence [6, p. 68] we have

$$[b+\gamma]^2[\text{meas } \Omega_*]\leqq \int_\Omega [Y\Phi]^2 \leqq \liminf_{\nu\to\infty}\int_\Omega [J_{\varepsilon_\nu}(U_{\varepsilon_\nu})\Phi]^2,$$

and, by an analogue of the Fatou lemma [6, p. 172], we have

$$\begin{aligned}\lim_{\nu\to\infty}\int_{\Omega_*}[J_{\varepsilon_\nu}(U_{\varepsilon_\nu})]^2 &= \limsup_{\nu\to\infty}\int_{\Omega_*}[J_{\varepsilon_\nu}(U_{\varepsilon_\nu})]^2\\ &\leqq \int_{\Omega_*}\limsup_{\nu\to\infty}[J_{\varepsilon_\nu}(U_{\varepsilon_\nu})]^2\\ &\leqq b^2[\text{meas } \Omega_*],\end{aligned}$$

since $\limsup_{\nu\to\infty}[J_{\varepsilon_\nu}(U_{\varepsilon_\nu})]^2\leqq b^2$ on Ω_*; this holds since U_{ε_ν} is convergent to 0 on Ω_*. This contradiction establishes (4.5) in the case of (4.6i). If (4.6ii) holds then

$$-\gamma[\text{meas } \Omega_*]\geqq \int_\Omega Y\Phi = \lim_{\nu\to\infty}\int_\Omega [J_{\varepsilon_\nu}(U_{\varepsilon_\nu})]\geqq 0$$

since $\liminf_{\nu\to\infty} J_{\varepsilon_\nu}(U_{\varepsilon_\nu})\geqq 0$ in Ω_*. This contradiction completely establishes (4.5).

Recall that, on the set Ω_0, we may define $H(U)(x)$ to be any value in $[0,b]$. We formally set

$$H(U)(x)=\begin{cases} Y(x), & x\in\Omega_0,\\ h(U(x)), & x\in\Omega-\Omega_0.\end{cases} \tag{4.7}$$

Now, by (4.2) and (4.4), we conclude that

$$Y=F_0-LV-G(U). \tag{4.8}$$

Thus, to establish the equation,

$$TV=LV+H(U)+G(U)=F_0 \tag{4.9}$$

it is necessary and sufficient to show that

$$Y(x)=h(U(x)), \qquad x\in\Omega-\Omega_0. \tag{4.10}$$

We shall verify (4.10) by proving that $J_{\varepsilon_\nu}(U_{\varepsilon_\nu})$ is weakly convergent in $L^2(\Omega-\Omega_0)$ to $H(U)$ which, by (4.4), must coincide with Y on $\Omega-\Omega_0$. Thus, defining

$$A_\nu=\{x\in\Omega-\Omega_0\colon U_{\varepsilon_\nu}(x)=0\},$$

we have, for fixed $\varphi\in L^2(\Omega-\Omega_0)$,

$$\begin{aligned}&\left|\int_{\Omega-\Omega_0}[H(U)-J_{\varepsilon_\nu}(U_{\varepsilon_\nu})]\varphi\right|\\ &\quad\leqq\left|\int_{A_\nu}[H(U)]\varphi\right|+\left|\int_{\Omega-\Omega_0-A_\nu}[H(U)-H(U_{\varepsilon_\nu})]\varphi\right|+\left|\int_{\Omega-\Omega_0}\tau_\nu\varphi\right|\end{aligned} \tag{4.11}$$

where

$$\tau_\nu(x)=\begin{cases}0 & \text{if } U_{\varepsilon_\nu}(x)=0,\\ h(U_{\varepsilon_\nu}(x))-j_{\varepsilon_\nu}(U_{\varepsilon_\nu}(x)) & \text{if } U_{\varepsilon_\nu}(x)\neq 0.\end{cases}$$

Now meas $(A_\nu)\to 0$ as $\nu\to\infty$ since U_ν converges in $L^2(\Omega)$ to U. In particular, the first term on the right side of (4.11) tends to zero. To see that $\{\tau_\nu\}$ is weakly convergent to zero in $L^2(\Omega-\Omega_0)$, note first that τ_ν converges pointwise to 0. If $x\in\Omega-\Omega_0$ is a point such that $U_{\varepsilon_\nu}(x)$ is convergent to $U(x)$, then $U_{\varepsilon_\nu}(x)\neq 0$ for all sufficiently large ν; say, $U_{\varepsilon_\nu}(x)$ lies in an open interval σ containing $U(x)$ and bounded away from 0 for $\nu\geqq\nu_0$. By Lemma 2.3, j_{ε_ν} converges uniformly on the compact set $\bar\sigma$ to h; in particular, $\tau_\nu(x)$ converges to zero. Thus, we may conclude from the Lebesgue dominated convergence theorem [6, p. 151] and the monotonicity of h and j_{ε_ν}, that the adjusted functions σ_ν,

$$\sigma_\nu(x)=\begin{cases}\tau_\nu(x), & |U_{\varepsilon_\nu}(x)-U(x)|\leqq 1,\\ 0 & \text{otherwise},\end{cases}$$

are convergent to 0 in $L^2(\Omega-\Omega_0)$. In particular, we obtain from this and from the inequality,

$$\left|\int_{\Omega-\Omega_0}\tau_\nu\varphi\right|\leqq\left|\int_{\Omega-\Omega_0}\sigma_\nu\varphi\right|+\left|\int_{B_\nu}[h(U_{\varepsilon_\nu})-j_{\varepsilon_\nu}(U_{\varepsilon_\nu})]\varphi\right|,$$

where

$$B_\nu=\{x\in\Omega-\Omega_0\colon|U_{\varepsilon_\nu}(x)-U(x)|>1\},$$

the required limit,

$$\lim_{\nu\to\infty}\int_{\Omega-\Omega_0}\tau_\nu\varphi=0, \tag{4.12}$$

since measure $(B_\nu)\to 0$ and since the sequence $h(U_{\varepsilon_\nu})-j_{\varepsilon_\nu}(U_{\varepsilon_\nu})$ is bounded in $L^2(\Omega)$. If we define $\{\varkappa_\nu\}\subset L^2(\Omega-\Omega_0)$ by

$$\varkappa_\nu(x)=\begin{cases}0 & \text{if } U_{\varepsilon_\nu}(x)=0,\\ h(U(x))-h(U_{\varepsilon_\nu}(x)) & \text{if } U_{\varepsilon_\nu}(x)\neq 0,\end{cases}$$

then the second term on the right side of (4.11) assumes the form $\int_{\Omega-\Omega_0}\varkappa_\nu\varphi$. An argument virtually indentical to that for $\{\tau_\nu\}$ shows that $\{\varkappa_\nu\}$ is weakly convergent to zero in $L^2(\Omega-\Omega_0)$, i.e.,

$$\lim_{\nu\to\infty}\int_{\Omega-\Omega_0}\varkappa_\nu\varphi=0. \tag{4.13}$$

Conditions (4.11), (4.12) and (4.13) now give the required weak convergence of $J_{\varepsilon_\nu}(U_{\varepsilon_\nu})$ to $H(U)$ in $L^2(\Omega-\Omega_0)$. In particular, $Y=H(U)$ and it follows that U is a solution of (2.14). The theorem follows from (4.1) and (4.4).

Remark. The referee has observed that a solution to the problem with $W=0$ and $f\in L^2(\Omega)$ can be obtained as follows. For fixed $v\in L^2(\Omega)$ define $u=T(v)\in H_0^1(\Omega)$ as the solution to $Lu+h(u)\ni f-g(v)$ (cf. Brezis–Strauss, J. Math. Soc. Japan, 25 (1973), pp. 565–590). A priori estimates of the form,

$$\|u\|_{H^1}\leqq C\|v\|_{L^2}+C\|f\|_{H^{-1}},$$

together with the compactness and continuity of T permit application of the Leray–Schauder degree theory and thus a fixed point for T.

Acknowledgment. The author expresses his gratitude to Haim Brezis, who generously furnished the proof of Theorem 2.2, and who suggested the separation of this existence result from the approximation result of Theorem 2.4; and to the referee, whose careful reading of the manuscript and subsequent recommendations effected considerable improvements.

REFERENCES

[1] S. AGMON, *Lectures on Elliptic Boundary Value Problems*, Van Nostrand, Princeton, NJ, 1965.

[2] J. P. AUBIN, *Approximation of Elliptic Boundary Value Problems*, Wiley-Interscience, New York, 1972.

[3] H. BREZIS, *Operateurs Maximaux Monotones et Semi-groupes de Contractions dans les Especes de Hilbert*, North-Holland, American Elsevier, Amsterdam and New York, 1973.

[4] ———, *Perturbation non linéaire d'opérateurs maximaux monotones*, C.R. Acad. Sci. Paris Sér. A-B, 269 (1969), pp. 566–569.

[5] JU. A. DUBINSKII, *Quasilinear elliptic and parabolic equations of arbitrary order*, Uspehi Mat. Nauk., 23 (1968), no. 1 (139), pp. 45–90.

[6] N. DUNFORD AND J. SCHWARTZ, *Linear Operators, Part I*, Wiley-Interscience, New York, 1957.

[7] J. JEROME, *Existence and approximation of weak solutions of the Stefan problem with nonmonotone nonlinearities*, Proceedings of the 1975 Dundee Conference on Numerical Analysis, Springer Lecture Notes in Mathematics 506, Springer-Verlag, Berlin, pp. 148–156.

[8] ———, *On the L_2 n-width of certain classes of functions of several variables*, J. Math. Anal. Appl., 20 (1967), pp. 110–123.

[9] ———, *Nonlinear equations of evolution and a generalized Stefan problem*, J. Differential Equations, 26 (1977), pp. 240–261.

[10] H. KELLER, *Elliptic boundary value problems suggested by nonlinear diffusion processes*, Arch. Rational Mech. Anal., 35 (1969), pp. 363–381.

[11] J. LIONS, *Quelques Méthodes de Résolution des Problèmes aux Limites non Linéaires*, Dunod, Paris, 1969.

[12] O. A. OLEINIK, *A method of solution of the general Stefan problem*, Soviet Math. Dokl., 1 (1960), pp. 1350–1354.

[13] S. PARTER, *Mildly nonlinear elliptic partial differential equations and their numerical solution*, Numer. Math., 7 (1965), pp. 113–128.

[14] M. ROSENZWEIG, *On weak solutions of a mildly nonlinear Dirichlet problem*, this Journal, 2 (1971), pp. 483–495.

[15] H. ROYDEN, *Real Analysis*, 2nd ed., Macmillan, Toronto, 1970.

[16] N. SCHRYER, *Newton's method for convex nonlinear elliptic boundary value problems*, Numer. Math., 17 (1971), pp. 284–300.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0016 $01.00/0

A STIELTJES INTEGRAL FORMULA FOR DANIELL FUNCTIONALS*

JAMES D. BAKER† AND JAMES A. DYER‡

Abstract. For the quasi-continuous function on a closed interval, the Daniell functionals are the left-Cauchy integrals which have nonincreasing, nonnegative, and left-continuous integrator functions. Relative to the dual space, the subspace generated by the Daniell functionals is shown to be a projection band and its orthogonal complement can be written as an interior integral with saltus integrator functions.

1. Introduction. A real valued function on $[a, b]$ is quasi-continuous if it is the uniform limit of a sequence of step-functions on $[a, b]$. This class of functions is a Banach space with the supremum norm and will be denoted by QC$[a, b]$.

A Daniell functional is defined to be a positive linear functional I on a vector lattice of functions with the property

$$\text{(D)} \qquad \lim_n I(f_n) = 0, \quad \text{where } f_1 \geqq f_2 \geqq \cdots, \quad \text{and} \quad \lim_n f_n = 0 \quad \text{pointwise.}$$

Since Stieltjes integrals are popular as representations for the bounded linear functionals on QC$[a, b]$, a similarity between these integrals and Daniell functionals is suggested. A relationship exists between Daniell functionals and measure integrals via Stone's theorem, and while Stieltjes integrals are not, in general, measure integrals, it is shown that the Daniell functionals can be characterized in terms of a special class of Stieltjes integrals. This analysis also provides a foundation for studying the relative structure of the subspace generated by the Daniell functionals with respect to the dual space of bounded linear functionals on QC$[a, b]$.

We conclude the introduction with an example of a non-Daniell functional. Let $[a, b]$ be a closed interval, $g(t) = 0$ if $a \leqq t < b$, and $g(b) = 1$. For $f \in$ QC$[a, b]$, the functional U defined by $U(f) = (L)\int_a^b f\,dg$ is positive and linear. The sequence of functions $f_n = \chi_{(b-1/n, b)}$ is nonincreasing with $\lim_n f_n = 0$ pointwise; however $\lim_n U(f_n) = 1$. Thus property (D) fails to hold.

2. Daniell functional representation. Our approach is to first show that QC$[a, b]$ is a Banach lattice ([6, pp. 224–236] and [7] can be used for background); to conclude from this that the linear functional defining the Daniell functional is a bounded linear functional on QC$[a, b]$; and to use a representation theorem for bounded linear functionals on QC$[a, b]$ to obtain the desired Stieltjes integral.

THEOREM 2.1. *The space* QC$[a, b]$ *is a Banach lattice.*

Proof. Most of the proof is straightforward, and we show only that the join defined by $f \vee g(x) = \max\{f(x), g(x)\}$ is quasi-continuous. Suppose $\varepsilon > 0$ and $t \in (a, b]$. If $f(t^-) = g(t^-)$, there is a $\delta > 0$ such that if $\zeta \in (t-\delta, t)$, then $|g(\zeta) - f(t^-)| < \varepsilon$ and $|f(\zeta) - f(t^-)| < \varepsilon$. Either $\max\{f(\zeta), g(\zeta)\} = f(\zeta)$ or $\max\{f(\zeta)\} = g(\zeta)$, and in either case, $|f \vee g(\zeta) - f(t^-)| < \varepsilon$. If $f(t^-) \neq g(t^-)$, let $k = \min\{\varepsilon, |f(t^-) - g(t^-)|\}$. There is a $\delta' > 0$ such that if $\zeta \in (t-\delta', t)$, then $|f(\zeta) - f(t^-)| < k/3$ and $|g(\zeta) - g(t^-)| < k/3$. Thus $\max\{f(\zeta), g(\zeta)\} = f(\zeta)$ if $f(t^-) > g(t^-)$ and $\max\{f(\zeta), g(\zeta)\} = g(\zeta)$ if $f(t^-) < g(t^-)$. In either case $|f \vee g(\zeta) - f \vee g(t^-)| < \varepsilon$. Since a similar argument holds for $f \vee g(t^+)$ for $t \in [a, b)$, we conclude that $f \vee g \in$ QC$[a, b]$.

It is known that a positive linear functional on a Banach lattice is a bounded linear functional [6, p. 239]; thus if I is a Daniell functional on QC$[a, b]$, then I is a bounded linear functional on QC$[a, b]$.

* Received by the editors December 11, 1975, and in revised form November 17, 1976.

† Honeywell Corporate Research Center, Bloomington, Minnesota 55420.

‡ Department of Mathematics, Iowa State University, Ames, Iowa 50010.

We use the representation theorem for bounded linear functionals on QC$[a, b]$ given in [1]. It states that if T is such a functional, then there are unique functions α and β of bounded variation on $[a, b]$ with $\alpha-\beta$ vanishing except at a countable number of points such that (see equation (1.1) of [1])

$$T(f)=f(a)\beta(a)+\int_a^b (\alpha, \beta)\, df. \tag{A}$$

The $\int_a^b (\alpha, \beta)\, df$ is the refinement limit of the approximating sums

$$\sum_{i-1}^{n} \{\alpha(t_{i-1})[f(\zeta_i)-f(t_{i-1})]+\beta(t_i)[f(t_i)-f(\zeta_i)]\}$$

where $P=\{t_0<t_1<\cdots<t_n\}$ is partition of $[a, b]$ and $\zeta_i \in (t_{i-1}, t_i)$ for $i=1, 2, \cdots, n$.

THEOREM 2.2. *If I is a Daniell functional on* QC$[a, b]$, *then there is a unique nonincreasing, left-continuous function β on $[a, b]$ with $\beta(b)\geqq 0$ such that*

$$I(f)=f(b)\beta(b)-(L)\int_a^b f\, d\beta. \tag{2.1}$$

Proof. Using property (D) and proceeding as in [2] with the basis functions given in [1], we have that $\alpha(t)=\beta(t^+)$ for $t\in[a, b)$ and that $\beta(t)=\beta(t^-)$ for $t\in(a, b]$. Thus,

$$I(f)=f(a)\beta(a)+\lim_P \sum_{i=1}^{n} \{\beta(t_{i-1}^+)[f(\zeta_i)-f(t_{i-1})]+\beta(t_i)[f(t_i)-f(\zeta_i)]\}. \tag{2.2}$$

Since β is left continuous, the limit term in (2.2) is $(R)\int_a^b \beta\, df$ when β is a step function. This is also true when β is a saltus function since f is bounded and β is the limit in variation of a sequence of left-continuous step-functions. Then equation (2.1) follows from the integration-by-parts theorem. When β is a continuous function of bounded variation, we have from Theorem 2.2 of [1] that

$$\int_a^b (\beta, \beta)\, df=f(b)\beta(b)-f(a)\beta(a)-(I)\int_a^b f\, d\beta$$

and for these conditions, the interior and left-Cauchy integrals are equivalent. Thus (2.1) holds when β is of bounded variation. The requirement that β be nonincreasing with $\beta(b)$ nonnegative is necessary to insure that I be positive.

To show that the representation in (2.1) characterizes a Daniell functional, we have the following result.

THEOREM 2.3. *Suppose γ is a nonincreasing, left-continuous function on $[a, b]$ with $\gamma(b)\geqq 0$. Then*

$$U(f)=f(b)\gamma(b)-(L)\int_a^b f\, d\gamma$$

defines a Daniell functional on QC$[a, b]$.

Proof. The observation that U is positive and linear follows directly from properties of the integral. Suppose $f_1\geqq f_2\geqq\cdots$ and $\lim_n f_n=0$ pointwise. For the case where β is a step-function, let $\{a=t_0<t_1\cdots t_n=b\}$ be a partition of $[a, b]$ which contains the points of discontinuity of β. If $\varepsilon>0$, choose N such that if $n\geqq N$, then $|f_n(t_i)-f(t_i)|<\varepsilon/V_a^b(\beta)$ for $i=0, 1, \cdots, n-1$. Then

$$\left|(L)\int_a^b f_n d\beta-(L)\int_a^b f\, d\beta\right|=\left|\sum_{i=1}^{n} [f_n(t_{i-1})-f(t_{i-1})][\beta(t_i)-\beta(t_{i-1})]\right|<\varepsilon.$$

For the case where β is a saltus function, let $\{\beta_m\}$ denote a sequence of left-continuous step functions which converge in variation to β. For each integer m, we have that $\lim_n (L)\int_a^b f_n\, d\beta_m = (L)\int_a^b f\, d\beta_m$ by the preceding argument. Since $|f_n| \leqq M$ for each integer n, it follows that

$$\left|(L)\int_a^b f_n\, d\beta_m - (L)\int_a^b f_n\, d\beta\right| \leqq M V_a^b(\beta_m - \beta),$$

and we have that $\lim_m (L)\int_a^b f_n\, d\beta_m = (L)\int_a^b f_n\, d\beta$ uniformly in n. Similarly, $\lim_n (L)\int_a^b f\, d\beta_m = (L)\int_a^b f\, d\beta$. Thus, by the iterated limits theorem,

$$\lim_n (L)\int_a^b f_n\, d\beta = (L)\int_a^b f\, d\beta.$$

For the case where β is a continuous function of bounded variation, the function f and each function f_n are measurable with respect to the measure $\mu(\beta)$ defined on a σ-algebra of subsets of $[a, b]$ containing open intervals and singleton points with the measure of an open interval defined to be the β-length. Then

$$(L)\int_a^b f_n\, d\beta = (LS)\int_{[a,b]} f_n\, d\mu(\beta),$$

$$(L)\int_a^b f\, d\beta = (LS)\int_{[a,b]} f\, d\mu(\beta),$$

and the theorem follows from the dominated convergence theorem for Lebesgue–Stieltjes integrals.

Theorems 2.2 and 2.3 along with Stone's theorem identify a class of Stieltjes integrals which have the properties of measure integrals. Also, analogous representations for the Daniell functional can be obtained in terms of right-Cauchy, interior, and Young integrals by employing the integration-by-parts formulas in [8].

3. The dual space. In this section we consider the relative subspace of QC*$[a, b]$ generated by the Daniell functionals. First we show that this subspace is a lattice ideal, then that it is a projection band, and finally, that its orthogonal complement can be expressed as a Stieltjes interior integral. For background on these topics the reader is referred to [7].

In studying the dual space, it is desirable to have an expression for the norm of the functionals. It is not clear that the representation in equation (A) can be used for this since

$$\|T\| \leqq |\beta(b)| + |\beta(a)| + V_a^b(\beta) + \sum_{t\in[a,b)} |\beta(t) - \alpha(t)|$$

and since there are functions α and β such that strict inequality holds. However, a functional representation introduced in [2], [3] is sufficient. We begin by reviewing some of these concepts and the reader is referred to [2] for details.

If X is a nonvoid set, a family, $\mathcal{P}$, of subsets of X is said to be a pre-algebra if the null set belongs to $\mathcal{P}$, if $\mathcal{P}$ is closed under intersection, and if differences of elements of $\mathcal{P}$ can be written as finite disjoint unions of sets in $\mathcal{P}$. A $\mathcal{P}$-volume is a finitely additive set function on $\mathcal{P}$. Q(X, $\mathcal{P}$) denotes the uniform closure of the scalar linear combination of characteristic functions of sets in $\mathcal{P}$. It is shown in [2] that if X is $[a, b]$ and $\mathcal{P}$ is the collection of open subintervals of $[a, b]$, singleton subsets of $[a, b]$, and ϕ then real

$Q(X, \mathscr{P})$ is $QC[a, b]$. A $\mathscr{P}$-subdivision of X is a disjoint cover of X by nonvoid elements of $\mathscr{P}$. The $\mathscr{P}$-subdivisions of X form a directed set when ordered by refinement. It is shown in [2] that if ϕ is a continuous linear functional on a $Q(X, \mathscr{P})$ space then there exists a unique $\mathscr{P}$-volume, μ, of bounded variation such that

$$\phi(f) = \psi \int_X f \, d\mu, \quad \forall f \in Q(X, \mathscr{P})$$

and

$$\mu(E) = \phi(\chi_E), \quad \forall E \in \mathscr{P}.$$

Here ψ denotes an arbitrary choice function of the nonvoid elements of $\mathscr{P}$, and the ψ-integral, $\psi \int_X f \, d\mu$ is defined as the limit under refinement of the approximating sums

$$\sum_{i=1}^{q} f(\psi(E_i))\mu(E_i),$$

where $\{E_i\}_{i=1}^{q}$ denotes an arbitrary $\mathscr{P}$-subdivision of X. Thus the adjoint of any $Q(X, \mathscr{P})$ may be identified with the family of $\mathscr{P}$-volumes of bounded variation normed with the variation norm. It is clear from the definition of the ψ-integral that if a real $Q(X, \mathscr{P})$ space is ordered in the same way that $QC[a, b]$ is ordered in this paper then a bounded linear functional is positive if and only if the associated $\mathscr{P}$-volume is positive. Throughout the remainder of this paper we will assume that X and $\mathscr{P}$ are such that $Q(X, \mathscr{P})$ is $QC[a, b]$ and we will identify $QC^*[a, b]$ with the space of $\mathscr{P}$-volumes of bounded variation. The Daniell functionals can be characterized in this representation in a fairly straightforward way.

THEOREM 3.1. *A $\mathscr{P}$-volume μ generates a Daniell functional on* $QC[a, b]$ *if and only if*

a) *μ is positive,*
b) $\lim_{t \to c-} \mu((t, c)) = 0, a < c \leqq b,$
c) $\lim_{t \to c+} \mu((c, t)) = 0, a \leqq c < b.$

Proof. Let ϕ denote the functional generated by μ. Then it follows that

$$\lim_{t \to c-} \mu((t, c)) = \lim_{t \to c-} \phi(\chi_{(t, c)}), \qquad a < c \leqq b,$$

and

$$\lim_{t \to c-} \chi_{(t,c)}(X) = 0, \qquad a < c \leqq b, \quad a \leqq X \leqq b.$$

Hence if ϕ is a Daniell functional it is easily seen that property b) holds. A similar argument shows that property c) holds and property a) follows from the previous discussion.

Conversely, suppose μ has properties a), b), and c). Define a real function g on $[a, b]$ by

$$g(t) = \begin{cases} \mu(\{t\}) + \mu((t, b)) + \mu(\{b\}), & a \leqq t < b, \\ \mu(\{b\}), & t = b. \end{cases} \tag{3.1}$$

Then g is nonnegative and if $a \leqq s < t \leqq b$ it follows that $g(s) - g(t)$ is equal to $\mu(\{s\}) + \mu((s, t))$, since μ is finitely additive. Thus g is nonincreasing and since for any $\mathscr{P}$-volume of bounded variation it must hold that

$$\lim_{s \to t-} \mu(\{s\}) = 0, \qquad a < t \leqq b,$$

it follows from property b) that g is left continuous. Thus by Theorem 2.3 the functional $\hat{\phi}$ defined by

$$\hat{\phi}(f)=f(b)g(b)-(L)\int_a^b f\,dg, \qquad f\in \mathrm{QC}[a,b]$$

is a Daniell functional. Let $\hat{\mu}$ be the associated $\mathscr{P}$-volume. Then it follows that

$$\hat{\mu}((c,d))=-(L)\int_a^b \chi_{(c,d)}\,dg=g(c+)-g(d), \qquad a\leqq c<d\leqq b.$$

Now it follows from property c), the properties of $\mathscr{P}$-volumes, and (3.1) that $g(c+)$ is $\mu(c,b)+\mu\{b\}$ and so we see that $\mu((c,d))$ is $\hat{\mu}(c,d)$. A similar argument shows that $\hat{\mu}(\{t\})$ is $\mu(\{t\})$, $a\leqq t\leqq b$ and so μ generates a Daniell functional. This completes the proof.

Note that this theorem also relates the ψ-integral for a $\mathscr{P}$-volume satisfying the conditions of the theorem to an ordinary left-Cauchy integral.

The characterization of Daniell functionals given in Theorem 3.1 makes the characterization of their structure as a subset of $\mathrm{QC}^*[a,b]$ a straightforward job. We assume from now on that $\mathrm{QC}^*[a,b]$ has been ordered with the canonical order induced by the given ordering on $\mathrm{QC}[a,b]$. We shall also identify $\mathrm{Q}^*[a,b]$ with its $\mathscr{P}$-volume representation whenever this is convenient.

Let D denote the linear subspace generated by the Daniell functionals. It is clear that D is simply the set of differences of such functionals. Thus D consists of those continuous linear functionals whose associated $\mathscr{P}$-volumes satisfy conditions b) and c) of Theorem 3.1. It is possible to show by a straightforward argument that if μ is a $\mathscr{P}$-volume of bounded variation then $|\mu|$ relative to the canonical order on $\mathrm{QC}^*[a,b]$ is the indefinite variation of μ. That is, $|\mu|$ is defined by

$$|\mu|(E)=\sup\sum_{i=1}^{q}|\mu(E_i)|, \quad \forall E\in\mathscr{P},$$

where the sup is taken over all $\mathscr{P}$-subdivision $\{E_i\}_{i=1}^q$ of E. By an argument almost the same as that used by Hildebrandt to prove Theorem II.4.7 in [4] one can show that

$$\lim_{t\to c-}[|\mu(t,c)|-|\mu|(t,c)]=0, \qquad a<c\leqq b,$$

and

$$\lim_{t\to c+}[|\mu(c,t)|-|\mu|(t,c)]=0, \qquad a\leqq c<b.$$

It then follows that if μ is in D so is $|\mu|$. Also in view of the form of conditions b) and c) of Theorem 3.1 it is easy to see that if ν is a $\mathscr{P}$-volume of bounded variation and for some μ in D, $|\nu|\leqq|\mu|$; then ν is in D. Consequently D is a lattice ideal.

Now suppose $\{\mu_n\}_{n=1}^\infty$ is a norm-convergent sequence in D converging to a $\mathscr{P}$-volume of bounded variation, μ. Let $\{\phi_n\}_{n=1}^\infty$ and ϕ be the associated linear functionals. Then if $\varepsilon>0$, and $a<c\leqq b$, there exists an n such that if $t<c$

$$\begin{aligned}|\mu((t,c))-\mu_n((t,c))|&=|\phi(\chi_{(t,c)})-\phi_n(\chi_{(t,c)})|\\&\leqq\|\phi_n-\phi\|<\varepsilon,\end{aligned}$$

since $\|\chi_{(t,c)}\|=1$. Consequently it follows that

$$\lim_{t\to c-}\mu(t,c)=0, \qquad a<c\leqq b.$$

A similar argument gives the same result for right hand limits in $[a, b)$. Thus μ is in D and D is closed. Since $\mathrm{QC}[a, b]$ is order complete we have shown the following theorem.

THEOREM 3.2. *Let D denote the set of all $\mathscr{P}$-volumes of bounded variation such that*

a) $\lim_{t\to c-} \mu((t, c)) = 0,\ a < c \leqq b,$

b) $\lim_{t\to c+} \mu((t, c)) = 0,\ a \leqq c < b.$

Then D is a band in $\mathrm{QC}^*[a, b]$.

This result together with Theorem 2.3 gives an alternative form for the theorem.

THEOREM 3.2′. *Let D denote the set of all linear functionals ϕ on $\mathrm{QC}[a, b]$ such that there exists a left continuous function g of bounded variation on $[a, b]$ such that*

$$\phi(f) = f(b)g(b) - (L)\int_a^b f\,dg, \quad \forall f \in \mathrm{QC}[a, b].$$

Then D is a band in $\mathrm{QC}^*[a, b]$.

It is also possible to characterize the orthogonal complement of D in terms of the $\mathscr{P}$-volume representation for $\mathrm{QC}^*[a, b]$. Following Schaefer's notation [7] we will denote this complement by $D^{\perp}$. A preliminary definition is useful for this characterization.

DEFINITION 3.1. If $a \leqq t < b$ then μ^t_{BR} will denote the $\mathscr{P}$-volume defined by

$$\mu^t_{\mathrm{BR}}(\{s\}) = 0, \qquad a \leqq s \leqq b,$$

$$\mu^t_{\mathrm{BR}}((c, d)) = \begin{cases} 1, & t \in [c, d), \\ 0 & \text{otherwise} \end{cases} \qquad a \leqq c < d \leqq b.$$

Similarly if $a < t \leqq b$ then μ^t_{BL} will denote the $\mathscr{P}$-volume defined by

$$\mu^t_{\mathrm{BL}}(\{s\}) = 0, \qquad a \leqq s \leqq b,$$

$$\mu^t_{\mathrm{BL}}((c, d)) = \begin{cases} 1, & t \in (c, d], \\ 0 & \text{otherwise} \end{cases} \qquad a \leqq c < d \leqq b.$$

These $\mathscr{P}$-volumes will be referred to as $\mathscr{P}$-volumes of type B.

It is easy to show that if f is in $\mathrm{QC}[a, b]$ then

$$\psi\int_{[a,b]} f\,d\mu^t_{\mathrm{BL}} = f(t-), \qquad a < t \leqq b,$$

and

$$\psi\int_{[a,b]} f\,d\mu^t_{\mathrm{BR}} = f(t+), \qquad a \leqq t < b.$$

Now if μ is any Daniell $\mathscr{P}$-volume and t is in $(a, b]$, then $\mu \wedge \mu^t_{\mathrm{BL}}(\{s\})$ is zero, $a \leqq t \leqq b$, and $\mu \wedge \mu^t_{\mathrm{BL}}((c, d))$ is zero if t is not in $(c, d]$. Let (c, d) be a subinterval of (a, b) such that t is in $(c, d]$. Then since $\mu \wedge \mu^t_{\mathrm{BL}}$ is finitely additive then

$$\mu \wedge \mu^t_{\mathrm{BL}}((a, b)) = \begin{cases} \mu \wedge \mu^t_{\mathrm{BL}}((c, d)), & t = d, \\ \mu \wedge \mu^t_{\mathrm{BL}}((c, t)), & t \neq d, \end{cases}$$

but since $\mu \wedge \mu^t_{\mathrm{BL}}((t, d))$ is zero we have that in any case $\mu \wedge \mu^t_{\mathrm{BL}}((a, b))$ equals $\mu \wedge \mu^t_{\mathrm{BL}}((c, d))$. Thus it follows that $\mu \wedge \mu^t_{\mathrm{BL}} = \alpha\mu^t_{\mathrm{BL}}$ for some $\alpha \geqq 0$. But $\mu \wedge \mu^t_{\mathrm{BL}}$ is in D since D is an ideal and so it follows that α is zero. Thus μ^t_{BL} is in $D^{\perp}$. A similar argument shows that μ^t_{BR} is in $D^\perp$, $a \leqq t < b$. In view of this result and in view of the functionals generated by $\mathscr{P}$-volumes of type B it is interesting to note that the $\mathscr{P}$-volume μ^t associated with evaluation at t is in D. The $\mathscr{P}$-volume μ^t is given by

$$\mu^t(\{s\}) = \delta_{ts}, \qquad a \leqq t \leqq b, \quad a \leqq s \leqq b,$$

$$\mu^t((c,d)) = \begin{cases} 1, & t \in (c,d), \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that μ is a nonnegative $\mathscr{P}$-volume. Throughout the remainder of this discussion if $a < t \leqq b$ we will denote $\lim_{s\to t-} \mu((s,t))$ by $\mu((t-,t))$ and if $a \leqq t < b$ we will denote $\lim_{s\to t+} \mu((t,s))$ by $\mu((t,t+))$. These limits necessarily exist since the nets involved are decreasing and bounded below by zero. Let $S = \{t_i\}_{i=1}^p$ be an arbitrary point subdivision of $[a, b]$. We will denote the family of all point subdivisions of $[a, b]$ by $\mathscr{D}$ and assume that the family is ordered by refinement. Define μ_S by the relation

$$\mu_S = \mu((a,a+))\mu^a_{\mathrm{BR}} + \mu((b-,b))\mu^b_{\mathrm{BL}} + \sum_{i=2}^{p-1} [\mu((t_i-,t_i))\mu^{t_i}_{\mathrm{BL}} + \mu((t_i,t_i+))\mu^{t_i}_{\mathrm{BR}}].$$

The net $\{\mu_S\}_{S\in\mathscr{D}}$ is an increasing net and since for any μ_S, $\|\mu_S\| = \mu_S((a,b))$ it follows that if S' refines S, $\|\mu_{S'} - \mu_S\| = |\mu_{S'}(a,b) - \mu_S(a,b)|$. Also for any S it is easily seen that

$$\mu_S((a,b)) = \mu((a,a+)) + \mu((b-,b)) + \sum_{i=2}^{p-1} [\mu((t_i-,t_i)) + \mu((t_i,t_i+))] \leqq \mu((a,b)).$$

Thus the net $\{\mu_S((a,b))\}_{S\in\mathscr{D}}$ is an increasing net of real numbers bounded above and hence converges. Therefore $\{\mu_S\}_{S\in\mathscr{D}}$ is a norm Cauchy net converging to a $\mathscr{P}$-volume μ_1 in D. Now suppose $a < c \leqq b$ and S be a point subdivision to which c belongs. If t is between c and the predecessor of c in S then $\mu_S((t,c))$ is equal to $\mu((c-,c))$, and thus

$$\lim_{t\to c-} \mu_S((t,c)) = \mu((c-,c)).$$

Therefore, since $\{\mu_S\}_{S\in\mathscr{D}}$ converges to μ in norm it follows that

$$\lim_{t\to c-} \mu_1((t,c)) = \mu((c-,c)),$$

and consequently that

$$\lim_{t\to c-} (\mu - \mu_1)((t,c)) = 0, \qquad a < c \leqq b.$$

A similar argument shows that

$$\lim_{t\to c+} (\mu - \mu_1)((t,c)) = 0, \qquad a \leqq c < b,$$

and therefore we see that $\mu - \mu_1$ is in D. This implies the following result.

THEOREM 3.3. *The orthogonal complement of D is the closed linear subspace generated by the $\mathscr{P}$-volumes of type* B.

The final theorem provides a representation for the functionals in D.

THEOREM 3.4. *If $\sigma \in D^\perp$, then there exists a saltus function s on $[a, b]$ such that*

$$\sigma(t) = f(b)s(b) - (I)\int_a^b f\,ds,$$

for $f \in \mathrm{QC}[a, b]$.

Proof. Let σ_R^t and σ_L^t denote the elements in QC*$[a, b]$ which correspond to μ_{BR}^t, and μ_{BL}^t respectively. For $t \in [a, b]$, let $\lambda_t = \chi_{[a,t]}$ and $\rho_t = \chi_{[a,t)}$. Using equation (A) with $\alpha = \beta = \rho_t$ along with Theorem 2.2 in [1], we have that $\sigma_R^t(f) = f(b)\rho_t(b) - (I)\int_a^b f\, d\rho_t$ for $f \in QC[a, b]$. Similarly, $\sigma_L^t(f) = f(b)\lambda_t(b) - (I)\int_a^b f\, d\lambda_t$. Since the elements in $D^\perp$ are limits in variation of a sequence made from linear combinations of the functional σ_R^t and σ_L^t, the theorem follows by letting s be the function corresponding to σ.

Acknowledgment. We are pleased to acknowledge suggestions of the referee which have led to many of the results presented in § 3.

REFERENCES

[1] J. D. BAKER, *Representation of linear functionals on quasi-continuous functions*, Proc. Amer. Math. Soc., 48 (1975), pp. 120–124.

[2] J. A. DYER, *Integral bases in linear topological spaces*, Illinois J. Math., 14 (1970), pp. 352–372.

[3] ——— *The Fredholm and Volterra problems for Stieltjes integral equations*, Applicable Anal., 5 (1975), 125–139.

[4] T. H. HILDEBRANDT, *Introduction to the Theory of Integration*, Academic Press, New York, 1963.

[5] ——— *Linear continuous functionals on the space (BV) with weak topologies*, Proc. Amer. Math. Soc., 17 (1966), pp. 658–664.

[6] J. L. KELLY AND I. NAMIOKA, *Linear Topological Spaces*, Van Nostrand, New York, 1963.

[7] H. H. SCHAEFER, *Banach Lattices and Positive Operators*, Springer-Verlag, Berlin, 1974.

[8] F. M. WRIGHT AND J. D. BAKER, *On integration-by-parts for weighted integrals*, Proc. Amer. Math. Soc., 22 (1969), pp. 42–52.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0017 $01.00/0

THE LAURENT EXPANSION OF A GENERALIZED RESOLVENT WITH SOME APPLICATIONS*

NICHOLAS J. ROSE†

Abstract. Let A, B be $n\times n$ complex matrices and assume $(A+\lambda B)^{-1}$ exists for some complex number λ; then $(A+\lambda B)^{-1}$ has a Laurent expansion of the form $\sum_{k=-\nu}^{\infty} Q_k\lambda^k$ with $Q_{-\nu}\neq 0$ valid in some deleted neighborhood of $\lambda=0$. Explicit formulas for the Q_k are given. Expansions of powers of $(A+\lambda B)^{-1}A$ are also given. The expansions are used to obtain various characterizations of the Drazin inverse and other inverses as limits. Finally the expansions, together with Laplace transforms, are used to solve the differential equations $A\dot{x}+Bx=0$ and $A\ddot{x}+Bx=0$, where A may be singular, in the case when unique solutions exist for appropriate initial conditions.

1. Introduction. Let $\mathbb{C}$ be the set of complex numbers and $\mathbb{C}^{m\times n}$ the set of $m\times n$ matrices with elements in $\mathbb{C}$. If $A, B\in\mathbb{C}^{n\times n}$ and $A+\lambda B$ is invertible for some $\lambda\in\mathbb{C}$, then the elements of $(A+\lambda B)^{-1}$ are rational functions of λ. Thus, for some $r>0$ and $0<|\lambda|<r$, we have the Laurent expansion

$$(1)\qquad (A+\lambda B)^{-1}=\sum_{k=-\nu}^{\infty} Q_k\lambda^k,\qquad Q_{-\nu}\neq 0,$$

where the coefficient matrices $Q_k\in\mathbb{C}^{n\times n}$ are independent of λ and are uniquely determined by A and B. The nonnegative integer ν is also independent of λ and uniquely determined by A and B. If $\nu>0$ then $(A+\lambda B)^{-1}$ has a pole of order ν at $\lambda=0$. When $B=I$, the identity matrix, $(A+\lambda I)^{-1}$, which always exists in a deleted neighborhood of $\lambda=0$, is often called the resolvent of A; thus $(A+\lambda B)^{-1}$ can be considered a generalized resolvent.

In [5] Langenhop has characterized the Q_k and ν; however, explicit representations were not given. In § 2 we present explicit expressions for the Q_k and ν as functions of A and B. In addition explicit expansions for powers of $(A+\lambda B)^{-1}A$ are given.

Various limits involving the resolvent $(A+\lambda I)^{-1}$ have been used by Ben-Israel [1], Meyer [7], Langenhop [5], to characterize the index of A and various pseudoinverses of A. In § 3 we obtain new limit theorems which generalize these results.

In § 4 we consider the system of differential equations $A\dot{x}+Bx=0$ where A may be a singular matrix. Using Laplace transforms and the Laurent expansion of the generalized resolvent we shall derive the solution in the case when unique solutions exist for appropriate initial conditions. This will reproduce some of the results in [3]. Finally in § 5, we solve the second order system $A\ddot{x}+Bx=0$ where A may be singular, in the case where unique solutions exist for appropriate initial conditions.

Our main tool will be the Drazin inverse of a square matrix. We briefly review some relevant definitions and facts. Further information may be found in [2, pp. 169–180], [3].

If $A\in\mathbb{C}^{n\times n}$ then the index of A, denoted by $\operatorname{ind} A$, is the least nonnegative integer ν such that rank $A^{\nu}=$ rank $A^{\nu+1}$ (we assume $A^0=I$ for all A). The Drazin inverse of $A\in\mathbb{C}^{n\times n}$, denoted by A^D, is the unique matrix X satisfying (i) $XA=AX$, (ii) $XAX=X$, (iii) $A^{k+1}X=A^k$ for $k\geqq \operatorname{ind} A$. In case $\operatorname{ind} A\leqq 1$, A^D is called the group inverse of A and denoted by $A^{\#}$. If $\operatorname{ind} A=\nu$ then there is a nonsingular matrix

* Received by the editors July 8, 1976, and in final revised form October 14, 1977.
† Department of Mathematics, North Carolina State University, Raleigh, North Carolina 27607.

$T \in \mathbb{C}^{n\times n}$ such that

$$A = T^{-1}\begin{bmatrix} C & 0 \\ 0 & N \end{bmatrix} T \tag{2}$$

where C is a nonsingular matrix and N is the nilpotent of index ν. (If $\nu = 0$, the N block is missing and if A is nilpotent the C block is missing.) If A is represented as in (2) then

$$A^D = T^{-1}\begin{bmatrix} C^{-1} & 0 \\ 0 & 0 \end{bmatrix} T. \tag{3}$$

The following facts about Drazin inverses will be needed:

1) A^D is a polynomial in A.
2) AA^D and $I - AA^D$ are idempotent.
3) If ind $A = \nu > 0$, then $(I - AA^D)A^k = 0$ for $k \geqq \nu$ but $(I - AA^D)A^{\nu-1} \neq 0$.
4) If $A, B \in \mathbb{C}^{n\times n}$, $AB = BA$ and $(A + \lambda B)^{-1}$ exists for some $\lambda \in \mathbb{C}$ then $(I - AA^D)BB^D = I - AA^D$.

The last result is in Lemma 1 of [3].

2. The expansion theorems. It is convenient to begin with the case when A and B commute.

THEOREM 2.1. *Assume $A, B \in \mathbb{C}^{n\times n}$, $AB = BA$ and $(A + \lambda B)^{-1}$ exists for some $\lambda \in \mathbb{C}$; then there exists an $r > 0$ such that the following hold for $0 < |\lambda| < r$:*

$$(A + \lambda B)^{-1} = A^D(I + \lambda A^D B)^{-1} + \lambda^{-1}B^D(I - AA^D)(I + \lambda^{-1}AB^D)^{-1}, \tag{4}$$

$$(A + \lambda B)^{-1} = A^D \sum_{k=0}^{\infty} (-1)^k (A^D B)^k \lambda^k + B^D(I - AA^D) \sum_{k=0}^{\nu-1} (-1)^k (AB^D)^k \lambda^{-k-1} \tag{5}$$

where $\nu =$ ind A. (If $\nu = 0$, the second term on the right of (5) is taken to be 0.)

Proof. Since $AB = BA$, the matrices A, A^D, B, B^D all commute. To prove (4) we start with the identity

$$(A + \lambda B)^{-1} = (A + \lambda B)^{-1}AA^D + (A + \lambda B)^{-1}(I - AA^D). \tag{6}$$

A direct calculation shows that the terms on the right of (6) are equal, respectively, to the terms on the right of (4).

Since $(I + \lambda A^D B)^{-1}$ can be expanded in a geometric series (Neumann series) in some neighborhood of $\lambda = 0$ we have

$$A^D(I - \lambda A^D B)^{-1} = A^D \sum_{k=0}^{\infty} (-1)^k (A^D B)^k \lambda^k$$

which is the first term on the right of (5). The second term of (4) involves $(I + \lambda^{-1}AB^D)^{-1}$ which exists in a deleted neighborhood of $\lambda = 0$, but which cannot be expanded in an infinite series in powers of $\lambda^{-1}AB^D$. However, we can use a finite geometric series plus a remainder

$$(I + \lambda^{-1}AB^D)^{-1} = \sum_{k=0}^{\nu-1} (-1)^k (AB^D)^k \lambda^{-k} + (-1)^\nu \lambda^{-\nu}(AB^D)^\nu (I + \lambda^{-1}AB^D)^{-1} \tag{7}$$

where we take $\nu =$ ind A (we assume $\nu = 0$, otherwise A is nonsingular and the second term in (5) is zero). Multiplying by $\lambda^{-1}B^D(I - AA^D)$ we find

$$\lambda^{-1}B^D(I - AA^D)(I + \lambda^{-1}AB^D)^{-1} = B^D(I - AA^D) \sum_{k=0}^{\nu-1} (-1)^k (AB^D)^k \lambda^{-k-1}$$

since $(I-AA^D)A^\nu=0$. This is the second term of (5). If $\nu>0$, the coefficient of $\lambda^{-\nu}$ is $(-1)^{\nu-1}(I-AA^D)A^{\nu-1}(B^D)^\nu$. It follows from the proof of Lemma 1 in [3] that this matrix is not 0 and thus $(A+\lambda B)^{-1}$ has a pole of order ν at $\lambda=0$.

We now consider the case when A and B do not necessarily commute. Let $\lambda_0\in\mathbb{C}$ be such that $(A+\lambda_0 B)^{-1}$ exists and let

$$\hat{A}=(A+\lambda_0B)^{-1}A,\qquad \hat{B}=(A+\lambda_0B)^{-1}B. \tag{8}$$

The matrices $\hat{A}$ and $\hat{B}$ depend on λ_0 and since $\hat{A}+\lambda_0\hat{B}=I$, $\hat{A}$ and $\hat{B}$ commute.

THEOREM 2.2. *Assume $A, B\in\mathbb{C}^{n\times n}$ and $(A+\lambda_0B)^{-1}$ exists for $\lambda_0\in\mathbb{C}$; then $(A+\lambda B)^{-1}$ exists in a deleted neighborhood of $\lambda=0$, and in this deleted neighborhood the following hold*:

$$(A+\lambda B)^{-1}=\{\hat{A}^D(I+\lambda\hat{A}^D\hat{B})^{-1}+\lambda^{-1}\hat{B}^D(I-\hat{A}\hat{A}^D)(I+\lambda^{-1}\hat{A}\hat{B}^D)^{-1}\}\cdot(A+\lambda_0B)^{-1}, \tag{9}$$

$$(A+\lambda B)^{-1}=\left\{\hat{A}^D\sum_{k=0}^{\infty}(-1)^k(\hat{A}^D\hat{B})^k\lambda^k+\hat{B}^D(I-\hat{A}\hat{A}^D)\sum_{k=0}^{\nu-1}(-1)^k(\hat{A}\hat{B}^D)^k\lambda^{-k-1}\right\}\cdot(A+\lambda_0B)^{-1} \tag{10}$$

where $\nu=\operatorname{ind}\hat{A}$ and $\hat{A}$, $\hat{B}$ are defined in (8). (*If $\nu=0$, the second term on the right of* (10) *is taken to be zero.*)

Proof. We rewrite $(A+\lambda B)^{-1}$ as follows:

$$\begin{aligned}(A+\lambda B)^{-1}&=(A+\lambda B)^{-1}(A+\lambda_0B)(A+\lambda_0B)^{-1}\\&=\{(A+\lambda_0B)^{-1}A+\lambda(A+\lambda_0B)^{-1}B\}^{-1}(A+\lambda_0B)^{-1}\\&=(\hat{A}+\lambda\hat{B})^{-1}(A+\lambda_0B)^{-1}.\end{aligned}$$

Since $\hat{A}$ and $\hat{B}$ commute, Theorem 2.1 may be applied to $(\hat{A}+\lambda\hat{B})^{-1}$ to yield (9) and (10). □

The coefficients of the various powers of λ in (10) appear to depend on λ_0. However, since the coefficient matrices are uniquely determined, different choices of admissible values of λ_0 must yield the same coefficient matrices. A direct proof of the fact that $\hat{A}^D\hat{B}$, $\hat{A}\hat{B}^D$, $\hat{A}^D(A+\lambda_0B)^{-1}$, $\hat{B}^D(A+\lambda_0B)^{-1}$ and $\operatorname{ind}\hat{A}$ are independent of λ_0 is given in [3].

For use in solving differential equations by Laplace transforms we need an expansion which holds for large values of the parameter. Letting $\lambda=z^{-1}$ in Theorem 2.2, we can easily prove:

COROLLARY 2.2. *Assume $A, B\in\mathbb{C}^{n\times n}$ and $(z_0A+B)^{-1}$ exists for $z_0\in\mathbb{C}$; then $(zA+B)^{-1}$ exists for $|z|>R$ for some $R>0$ and the following hold for $|z|>R$*:

$$(zA+B)^{-1}=(\tilde{A}^D(zI+\tilde{A}^D\tilde{B})^{-1}+\tilde{B}^D(I-\tilde{A}\tilde{A}^D)(I+z\tilde{A}\tilde{B}^D)^{-1}\}(z_0A+B)^{-1}, \tag{11}$$

$$(zA+B)^{-1}=\left\{\tilde{A}^D\sum_{k=0}^{\infty}(-1)^k(\tilde{A}^D\tilde{B})^kz^{-k-1}+\tilde{B}^D(I-\tilde{A}\tilde{A}^D)^kz^k\right\}(z_0A+B)^{-1} \tag{12}$$

where $\tilde{A}=(z_0A+B)^{-1}A$, $\tilde{B}=(z_0A+B)^{-1}B$ and $\nu=\operatorname{ind}\tilde{A}$. (*The second term of* (12) *is zero if $\nu=0$.*)

Returning to the case when A and B commute, we see that $(A+\lambda B)^{-1}$ exists then $(A+\lambda B)^{-l}$ exists for positive integral l. A Laurent expansion is provided in the following theorem.

THEOREM 2.3. *If $A, B \in \mathbb{C}^{n\times n}$, $AB = BA$ and $(A+\lambda B)^{-1}$ exists for some $\lambda \in \mathbb{C}$, then in some deleted neighborhood of $\lambda = 0$ the following hold*:

(13) $(A+\lambda B)^{-l} = (A^D)^l(I+\lambda A^D B)^{-l} + (B^D)^l(I-AA^D)(\lambda I + AB^D)^{-l}$,

(14) $(A+\lambda B)^{-l} = (A^D)^l \sum_{k=0}^{\infty} \binom{-l}{k}(A^D B)^k\lambda^k + (B^D)^l(I-AA^D)\sum_{k=0}^{\nu-1}\binom{-l}{k}(AB^D)^k\lambda^{-k-1}$

where $\nu = \operatorname{ind} A$, l is a positive integer and $\binom{-l}{k}$ are the binomial coefficients. (If $\nu = 0$ the second term in (14) *is taken to be zero.)*

Proof. The proof follows the same lines as Theorem 2.1 except that the binomial theorem is used in place of the geometric series. □

If A and B do not commute, we note that the expansion of $(A+\lambda B)^{-1}$ in (10) contains the term $(A+\lambda_0 B)^{-1}$ on the right so that all the coefficients do not commute. However, $(A+\lambda B)^{-1}A = (\hat{A}+\lambda\hat{B})^{-1}\hat{A}$ and all the terms commute. We therefore can write a formula for $((A+\lambda B)^{-1}A)^l$ where l is a positive integer.

THEOREM 2.4. *If $A, B \in \mathbb{C}^{n\times n}$ and $(A+\lambda_0 B)^{-1}$ exists for $\lambda_0 \in \mathbb{C}$, then, for positive integers l, $(A+\lambda B)^{-l}$ exists in a deleted neighborhood of $\lambda = 0$ and the following expansion holds*:

(15) $((A+\lambda B)^{-1}A)^l = \hat{A}\hat{A}^D \sum_{k=0}^{\infty}\binom{-l}{k}(\hat{A}^D\hat{B})^k\lambda^k + (I-\hat{A}\hat{A}^D)\sum_{k=0}^{\nu-1}\binom{-l}{k}(\hat{A}\hat{B}^D)^{k+l}\lambda^{-k-1}$

where $\nu = \operatorname{ind}\hat{A}$ and if $\nu = 0$ or $l \geqq \nu$, the second term on the right of (15) *is zero.*

Proof. $((A+\lambda B)^{-1}A)^l = ((\hat{A}+\lambda\hat{B})^{-1}\hat{A})^l = (\hat{A}+\lambda\hat{B})^{-l}\hat{A}^l$. We now apply Theorem 2.3 to obtain (15). □

3. Limit theorems. If $A \in \mathbb{C}^{n\times n}$ and l is a positive integer, $(A+\lambda I)^{-l}$ always exists in a deleted neighborhood of $\lambda = 0$ and we may set $B = I$ in (14) to obtain

$$(16)\qquad (A+\lambda I)^{-l} = (A^D)^l\sum_{k=0}^{\infty}\binom{-l}{k}(A^D)^k\lambda^k + (I+AA^D)\sum_{k=0}^{\nu-1}\binom{-l}{k}A^k\lambda^{-k-1}$$

where $\nu = \operatorname{ind} A$. It is easy to see that $(A+\lambda I)^{-l}$ is analytic at $\lambda = 0$ if and only if $\nu = 0$ (A^{-1} exists). If $\nu = 0$, we may make $(A+\lambda I)^{-l}$ analytic by multiplying by λ^ν or by multiplying by A^ν or by multiplying by $\lambda^m A^p$ where $m+p \geqq \nu$. Since $\lambda^m A^p(A+\lambda I)^{-l}$ is analytic at $\lambda = 0$ if $m+p \geqq \nu$, the limit of this expression will exist as $\lambda \to 0$. The following theorem is an easy consequence of these remarks and generalizes some of the limit theorems in [1], [6], [7].

THEOREM 3.1. *If $A \in \mathbb{C}^{n\times n}$ and $\lambda \in \mathbb{C}$ then*:

1. $\operatorname{ind} A \leqq 1$ *($A^{\#}$ exists) if and only if $(A+\lambda I)^{-2}A$ is analytic at $\lambda = 0$ and in this case*

$$\lim_{\lambda\to 0}(A+\lambda I)^{-2}A = A^{\#};$$

2. $\operatorname{ind} A$ *is the least nonnegative integer ν for which $(A+\lambda I)^{-l}A^\nu$ or $\lambda^\nu(A+\lambda I)^{-l}$ is analytic at $\lambda = 0$ ($l = 1, 2, \cdots$) in which case*

$$\lim_{\lambda\to 0}(A+\lambda I)^{-l}A^\nu = (A^D)^l A^\nu,$$

$$\lim_{\lambda\to 0}\lambda^\nu(A+\lambda I)^{-l} = \binom{-l}{\nu-1}(I-AA^D)A^{\nu-1};$$

3. *if* $l \geqq \operatorname{ind} A$ *then* $(A+\lambda I)^{-l-1}A^l$ *is analytic at* $\lambda = 0$ *and*

$$\lim_{\lambda\to 0}(A+\lambda I)^{-l-1}A^l = A^D.$$

THEOREM 3.2. *If* $A, B \in \mathbb{C}^{n\times n}$ *and* $(A+\lambda_0 B)^{-1}$ *exists for* $\lambda_0 \in \mathbb{C}$ *then* $(A+\lambda B)^{-1}A$ *is analytic at* $\lambda = 0$ *if and only if* $\operatorname{ind}\hat{A} \leqq 1$ $(\hat{A} = (A+\lambda_0 B)^{-1}A)$, *and in this case*

$$\lim_{\lambda\to 0}(A+\lambda B)^{-1}A = \hat{A}\hat{A}^D. \tag{17}$$

Proof. From equation (15) with $l=1$ we see that $(A+\lambda B)^{-1}A$ is analytic at $\lambda = 0$ if and only if $(I-\hat{A}\hat{A}^D)\hat{A}\hat{B}^D = 0$. Using the fact that $\hat{B} = (I-\hat{A})\lambda_0^{-1}$ (we may assume $\lambda_0 \neq 0$) and the representation for $\hat{A}$ given in (2), we find that $(I-\hat{A}\hat{A}^D)\hat{A}\hat{B}^D = 0$ if and only if $\operatorname{ind}\hat{A} \leqq 1$. The limit in (17) follows easily from (15). □

As a corollary of the above, we may present a limit suggested in Chipman [4] and discussed in Ward [9].

COROLLARY 3.2. *Let* $E \in \mathbb{C}^{p\times n}$, $F \in \mathbb{C}^{q\times n}$ *and assume that* $U = (E^*E+\lambda_0 F^*F)^{-1}$ *exists for some* $\lambda_0 > 0$; *then* $(E^*E+\lambda F^*F)^{-1}E^*E$ *and* $(E^*E+\lambda F^*F)^{-1}E^*$ *are analytic at* $\lambda = 0$ *and*

$$\lim_{\lambda\to 0}(E^*E+\lambda F^*F)^{-1}E^*E = (UE^*E)^D(UE^*E) \tag{18}$$

and

$$\lim_{\lambda\to 0}(E^*E+\lambda F^*F)^{-1}E^* = (UE^*E)^D UE^*. \tag{19}$$

Proof. Let $A = E^*E$ and $B = F^*F$ in (17); then $\hat{A} = UE^*E$. Equation (18) follows from (17) if we prove $\operatorname{ind}\hat{A} \leqq 1$. It is clear that U is Hermitian and positive definite. Assume $\hat{A}$ is singular and let $x \in \mathbb{C}^{n\times 1}$, $x \neq 0$, satisfy $\hat{A}^2 x = (UE^*E)^2 x = 0$. Then letting $\langle x, y\rangle$ denote the complex inner product for $x, y \in \mathbb{C}^{n\times 1}$ we have with $U^{1/2}$ the positive definite square root of U:

$$0 = \langle U^{-1}x, (UE^*E)^2 x\rangle = \langle U^{1/2}E^*Ex, U^{1/2}E^*Ex\rangle;$$

thus $U^{1/2}E^*Ex = 0$, $\hat{A}xUE^*Ex = 0$ and $\operatorname{ind}\hat{A} = 1$. To prove (19) we have, letting $E^\dagger$ denote the Moore–Penrose inverse of E,

$$\begin{aligned}\lim_{\lambda\to 0}(E^*E+\lambda F^*F)^{-1}E^* &= \lim_{\lambda\to 0}(E^*E+\lambda F^*F)^{-1}E^*EE^\dagger\\ &= (UE^*E)^D(UE^*E)E^\dagger\\ &= (UE^*E)^D(UE^*). \quad □\end{aligned}$$

If we let $F = I$ in (19) we get the well-known result [1]

$$\lim_{\lambda\to 0}(E^*E+\lambda I)^{-1}E^* = (E^*E)^D E^* = E^\dagger.$$

As is pointed out in Ward [9], the right-hand side of (19) can be considered a "weighted" generalized inverse of E.

THEOREM 3.3. *Let* $A, B \in \mathbb{C}^{n\times n}$ *and* $(A+\lambda_0 B)^{-1}$ *exists for* $\lambda_0 \in \mathbb{C}$ *and let* $\hat{A} = (A+\lambda_0 B)^{-1}A$. *If* l *is a positive integer and* $l \geqq \operatorname{ind}\hat{A}$ *then*

$$\lim_{\lambda\to 0}((A+\lambda B^{-1})A)^l = \hat{A}\hat{A}^D. \tag{20}$$

Proof. This follows directly from Theorem 2.4. □

We remark that equation (20) shows that $\hat{A}\hat{A}^D$ is independent of λ_0.

4. Applications to differential equations. Consider the initial value problem

$$A\dot{x}(t)+Bx(t)=f(t), \qquad 0\leqq t<\infty; \quad x(0)=x_0, \tag{21}$$

where $A, B\in\mathbb{C}^{n\times n}$ and $x(t), f(t)\in\mathbb{C}^{n\times 1}$ for all t. We allow A to be singular and assume f is sufficiently smooth. If a solution exists for a particular x_0, x_0 is said to be a *consistent initial condition.* In [3] it is shown that unique solutions exist for consistent initial conditions if and only if $(z_0A+B)^{-1}$ exists for some $z_0\in\mathbb{C}$. Characterizations of consistent initial conditions and an explicit formula for the solution are also given in [3]. In this section we indicate how the solution may be obtained by a Laplace transform method and the use of the Laurent expansion for the generalized resolvent. For brevity, we consider only the homogeneous differential equation

$$A\dot{x}+Bx=0, \qquad x(0)=x_0. \tag{22}$$

Assuming solutions exist and possess Laplace transforms, we take the Laplace transform of the differential equation to obtain

$$(zA+B)\mathscr{L}x=Ax_0$$

where $\mathscr{L}x$ is the Laplace transform of x.

The assumption that $(zA+B)^{-1}$ exists for some $z\in\mathbb{C}$ ensures that $(zA+B)^{-1}$ exists for $|z|$ sufficiently large. Thus

$$\mathscr{L}x=(zA+B)^{-1}Ax_0. \tag{23}$$

Using the expansion (12) we find

$$\mathscr{L}x=\left\{\tilde{A}\tilde{A}^D(zI+\tilde{A}^D\tilde{B})^{-1}+\tilde{A}\tilde{B}^D(1-\tilde{A}\tilde{A}^D)\sum_{k=0}^{\nu-1}(-1)^k(\tilde{A}\tilde{B}^D)^kz^k\right\}x_0. \tag{24}$$

The Laplace transform $\mathscr{L}x$ must approach zero as $\operatorname{Re}(z)\to\infty$ [8, p. 340]. This is the case if and only if

$$\tilde{A}\tilde{B}^D(I-\tilde{A}\tilde{A}^D)x_0=0. \tag{25}$$

If x_0 satisfies (25) we are left with

$$\mathscr{L}x=\tilde{A}\tilde{A}^D(zI+\tilde{A}^D\tilde{B})^{-1}x_0 \tag{26}$$

but, it is well-known that if $A\in\mathbb{C}^{n\times n}$, $\mathscr{L}(\exp At)=(zI-A)^{-1}$; therefore

$$x(t)=\tilde{A}\tilde{A}^D e^{-\tilde{A}^D\tilde{B}t}x_0. \tag{27}$$

However, since $x(t)$ must be continuous we must have $x(t)\to x_0$ as $t\to 0$; letting $t\to 0$ in (27) we find

$$x_0=\tilde{A}\tilde{A}^Dx_0 \quad\text{or}\quad (I-\tilde{A}\tilde{A}^D)x_0=0. \tag{28}$$

Note that (28) implies (25) so that consistent initial conditions are characterized by (28) and if x_0 satisfies (28), the unique solution is given by (27). □

If x_0 does not satisfy (28) then (27) still provides a solution of the differential equation but $x(t)$ will not approach x_0 as $t\to 0$.

5. Solution of $A\ddot{x}+Bx=0$. We now consider the homogeneous second order equation

$$A\ddot{x}(t)+Bx(t)=0, \qquad 0\leqq t<\infty, \tag{29}$$

together with the initial conditions

(30) $$x(0)=x_0, \qquad \dot{x}(o)=v_0,$$

where $A, B \in \mathbb{C}^{n\times n}$, A may be singular and $x(t)\in\mathbb{C}^{n\times 1}$. Assuming a solution exists and possesses a Laplace transform we find

(31) $$(z^2A+B)\mathcal{L}x = A(zx_0+v_0).$$

We assume that $(z_0^2A+B)^{-1}$ exists for $z_0\in\mathbb{C}$; this ensures that $(z^2A+B)^{-1}$ exists for $|z|$ sufficiently large. Thus

$$\mathcal{L}x=(z^2A+B)^{-1}A(zx_0+v_0).$$

Letting $\tilde{A}=(z_0^2A+B)^{-1}A$, $\tilde{B}=(z_0^2A+B)^{-1}B$, we find using (12) that

(32) $$\mathcal{L}x=(z^2I+\tilde{A}^D\tilde{B})^{-1}\tilde{A}^D\tilde{A}(zx_0+v_0)$$

providing x_0 satisfies

(33) $$(I-\tilde{A}\tilde{A}^D)\tilde{A}\tilde{B}^Dx_0=0 \quad\text{and}\quad (I-\tilde{A}\tilde{A}^D)\tilde{A}\tilde{B}^Dv_0=0.$$

From (32) we obtain

(34) $$\mathcal{L}x=\tilde{A}\tilde{A}^D\left(\frac{I}{z^2}-\frac{(\tilde{A}^D\tilde{B})}{z^4}+\frac{(\tilde{A}^D\tilde{B})^2}{z^6}+\cdots\right)(zx_0+v_0).$$

We may take inverse transforms of (34) term by term. It is convenient to define the functions $S(t)$, $C(t)$ by the everywhere convergent series:

(35) $$S(t)=\left(tI-\frac{\tilde{A}^D\tilde{B}t^3}{3!}+\frac{(\tilde{A}^D\tilde{B})^3t^5}{5!}-\cdots\right),$$

(36) $$C(t)=\left(I-\frac{(\tilde{A}^D\tilde{B})t^2}{2!}+\frac{(\tilde{A}^D\tilde{B})^2t^4}{4!}+\cdots\right);$$

then we find

(37) $$x(t)=\tilde{A}\tilde{A}^DC(t)x_0+\tilde{A}\tilde{A}^DS(t)v_0.$$

Letting $t\to 0$ in (37) and the derivative of (37) we find

(38) $$x_0=\tilde{A}\tilde{A}^Dx_0, \qquad v_0=\tilde{A}\tilde{A}^Dv_0.$$

If x_0, v_0 satisfy (38) then (33) is satisfied; thus equations (38) characterize consistent initial conditions and the solution is then given by (37).

Acknowledgments. The author would like to thank C. D. Meyer, Jr. and S. L. Campbell for many helpful discussions in the preparation of this paper. Professor Campbell suggested the proof of the first part of Corollary 3.2.

REFERENCES

[1] A. BEN-ISRAEL, *On matrices of index zero or one*, SIAM J. Appl. Math., 17 (1969), pp. 1118–1121.

[2] A. BEN-ISRAEL AND T. N. E. GREVILLE. *Generalized Inverses—Theory and Application*, Wiley–Interscience, New York, 1974.

[3] S. L. CAMPBELL, C. D. MEYER, JR. AND N. J. ROSE, *Applications of the Drazin inverse to linear systems of differential equations with singular constant coefficients*, SIAM J. Appl. Math., 31 (1976), pp. 411–425.

[4] J. S. CHIPMAN, *Least squares with insufficient observations*, J. Amer. Statist. Assoc., 54 (1964), pp. 1078–1111.

[5] C. E. LANGENHOP, *The Laurent expansion of a nearly singular matrix*, Linear Algebra and Appl., 4 (1971), pp. 329–340.
[6] ———, *On the index of a square matrix*, SIAM J. Appl. Math., 21 (1971), pp. 191–194.
[7] C. D. MEYER, JR., *Limits and the index of a square matrix*, Ibid., 26 (1974), pp. 469–478.
[8] F. D. MURNAGHAN, *Introduction to Applied Mathematics*, John Wiley, New York, 1948.
[9] J. F. WARD, JR., *On a limit formula for weighted pseudoinverses*, SIAM J. Appl. Math., 33 (1977), pp. 34–38.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0018 $01.00/0

MONOTONICITY OF THE ZEROS OF A CROSS-PRODUCT OF BESSEL FUNCTIONS*

MOURAD E. H. ISMAIL† AND MARTIN E. MULDOON‡

Abstract. Our principal result is that for fixed $\beta (0<\beta \leqq 1)$, and fixed $\alpha>0$, the positive zeros of the cross-product

$$J_{\nu+\beta}(x)K_\nu(\alpha x)-\alpha^\beta J_\nu(x)K_{\nu+\beta}(\alpha x)$$

increase with ν, $-\beta/2\leqq \nu<\infty$. In particular this implies that the eigenvalues of the boundary value problem

$$\nabla_n^2 p+\lambda^2 g(x)p=0,$$

p radial, $p'(0)=0$, $p(\infty)<\infty$, increase with the dimension n where ∇_n^2 is the n-dimensional Laplacian, x is the distance from the origin and $g(x)=1$, $0\leqq x\leqq 1$, $g(x)=-\alpha^2$, $x>1$.

1. Introduction. T. Nagylaki considered the boundary value problem given by the equation

$$\nabla_n^2 p+\lambda^2 g(x)p=0, \tag{1.1}$$

where p is radial, $p'(0)=0$, $p(\infty)<\infty$, ∇_n^2 is the n-dimensional Laplacian, x the distance from the origin and

$$g(x)=\begin{cases}1, & 0\leqq x\leqq 1,\\ -\alpha^2, & x>1.\end{cases} \tag{1.2}$$

(In case $n\geqq 2$, the condition $p'(0)=0$ may be replaced by $p(0)<\infty$) He conjectured that for each fixed $\alpha(>0)$ the smallest positive eigenvalue $\lambda_n(\alpha)$ increases with the dimension n, i.e., that

$$\lambda_n(\alpha)<\lambda_{n+1}(\alpha), \qquad n=1,2,\cdots, \qquad \alpha>0, \tag{1.3}$$

and showed [9, pp. 613–614] that $\lambda_1(\alpha)<\lambda_2(\alpha)$ for small α, for large α, and for $\alpha=\frac{1}{2}, 1, 2$ and mentioned the desirability of proving this for all positive α. Nagylaki was led to (1.3) in a study of the one- and two-dimensional cases in a study of migration and selection [9].

If the method of separation of variables is applied to (1.1) we are led, for the radial part, to the boundary value problem

$$(-xy')'+(\nu^2/x)y=\lambda^2 xg(x)y, \qquad p'(0)=0, \quad p(\infty)<\infty, \quad y(x)=x^\nu p(x), \tag{1.4}$$

where $\nu=(n-2)/2$. As we shall see in § 4, the positive eigenvalues are the positive zeros of fixed rank of the cross-product

$$J_{\nu+1}(x)K_\nu(\alpha x)-\alpha K_{\nu+1}(\alpha x)J_\nu(x) \tag{1.5}$$

with the usual notation [13] for Bessel functions and modified Bessel functions. It is natural, then, to generalize Nagylaki's conjecture to the assertion that for each fixed $\alpha>0$, the smallest positive zero of the cross-product (1.5) increases with ν, $\nu\geqq -\frac{1}{2}$. In

* Received by the editors May 14, 1976, and in revised form September 15, 1976.

† Department of Applied Mathematics, McMaster University, Hamilton, Ontario, Canada L8S 4L8. The work of this author was supported by the National Research Council of Canada under Grants A4048 and A3049.

‡ Department of Mathematics, York University, Downsview, Ontario, Canada M35 1P3. The work of this author was supported by the National Research Council of Canada under Grant A5199.

a private communication with us, R. Askey has in fact conjectured that for each $\alpha > 0$ and each fixed $\beta > 0$ the smallest positive zero of

$$J_{\nu+\beta}(x)K_\nu(\alpha x) - \alpha^\beta J_\nu(x)K_{\nu+\beta}(\alpha x) \tag{1.6}$$

increases with ν ($\nu \geqq 0$).

Our principal result (Theorem 3.2) is a generalization of Askey's conjecture in several ways in the case $0 < \beta \leqq 1$; it refers to a slightly more general cross-product, it is valid for $\nu \geqq -\beta/2$ and it asserts that a positive zero of fixed rank (first, second, etc.) increases with ν. In particular it implies Nagylaki's conjecture (1.3).

The proof of Theorem 3.2 depends on several preliminary results concerning the monotonicity of ratios of Bessel functions. These are given in § 2.

There is another approach to the problem of the monotonicity of eigenvalues of (1.4) based on a fairly standard argument of Sturmian type. This s outlined in § 4. This method is confined to the zeros of (1.5) rather than (1.6) and it gives results only for $\nu \geqq 0$ so that in particular it gives (1.3) only for $n = 2, 3, \cdots$. However it can be extended to boundary value problems with choices of $g(x)$ different from the one given by (1.2). The method also gives results on the monotonicity of $\lambda(\nu)/\nu$ where $\lambda(\nu)$ is an eigenvalue of (1.4). The precise result is given in Theorem 4.1.

One of our preliminary results on modified Bessel functions (Lemma 2.2) implies in particular that for all real ν

$$K_{\nu-1}(x)K_{\nu+1}(x) - K_\nu^2(x) > 0, \qquad x > 0. \tag{1.7}$$

This is an inequality of Turán type; see Szegö [11] where the Turán inequality

$$[P_n(x)]^2 - P_{n-1}(x)P_{n+1}(x) \geqq 0, \qquad n \geqq 1, \quad 0 \leqq x \leqq 1,$$

for Legendre polynomials is discussed in detail.

Lemma 2.2 implies also that $f(\nu, x)$ is a decreasing function of ν, where

$$f(\nu, x) = K_\nu(x)/K_{\nu+1}(x).$$

We believe, but are unable to prove, that $f(\nu, x)$ is a completely monotonic function of ν^2. In [6], Ismail and Kelker conjectured that $x^{-1/2}K_\nu(\sqrt{x})/K_{\nu+1}(\sqrt{x})$ is a completely monotonic function of x for $\nu \geqq 0$. Grosswald [4] proved this conjecture. Thus, if our present conjecture is correct, then the function $x^{-1/2}K_\nu(\sqrt{x})/K_{\nu+1}(\sqrt{x})$ will not only be completely monotonic in x for each fixed $\nu \geqq 0$ but also completely monotonic in ν^2 for each fixed positive x.

2. Monotonicity properties of ratios of Bessel functions. Most of the results in this section are based on Nicholson's formula [3, p. 54, (39)]

$$K_\mu(z)K_\nu(z) = 2\int_0^\infty K_{\mu+\nu}(2z \cosh t) \cosh[(\mu-\nu)t]\, dt, \qquad \operatorname{Re} z > 0, \tag{2.1}$$

and Neumann's formula [3, p. 47, (11)]

$$J_\mu(z)J_\nu(z) = (2/\pi)\int_0^{\pi/2} J_{\mu+\nu}(2z \cos\theta) \cos[(\mu-\nu)\theta]\, d\theta, \qquad \operatorname{Re}(\mu+\nu) > -1. \tag{2.2}$$

We use $j_{\nu k}$ to denote the kth positive zero of $J_\nu(x)$; we put $j_{\nu 0} = 0$.

We also need to use the fact, proved by R. G. Cooke [1], [2], and in a simpler way by J. Steinig [10], that

$$\int_0^x J_\nu(t)\, dt > 0, \qquad x > 0, \quad \nu > -1.$$

A simple application of the second mean value theorem then gives the following result.

LEMMA 2.1 *Let $\phi(t)$ be positive nonincreasing and continuous for $0<t<x$. Then, for $\nu>-1$ and $x>0$*

$$\int_0^x J_\nu(t)\phi(t)\,dt>0.$$

We now give a sequence of four lemmas on monotonicity of ratios of Bessel functions.

LEMMA 2.2. *For each fixed $x>0$ and each fixed $\beta>0$, the function $K_{\nu+\beta}(x)/K_\nu(x)$ increases with ν, $-\infty<\nu<\infty$.*

Proof. Since $K_\nu(x)$ is positive for positive x it suffices to show that, for each $\varepsilon>0$,

$$K_\nu(x)K_{\nu+\beta+\varepsilon}(x)-K_{\nu+\beta}(x)K_{\nu+\varepsilon}(x)>0. \tag{2.3}$$

Using (2.1), we find that the left hand side of (2.3) is equal to

$$2\int_0^\infty K_{2\nu+\varepsilon+\beta}(2x\cosh t)\{\cosh[(\beta+\varepsilon)t]-\cosh[(\beta-\varepsilon)t]\}\,dt$$

and this is positive since $\cosh x$ is an even function which is increasing for $x>0$ and $|\beta-\varepsilon|<\beta+\varepsilon$ for $\varepsilon>0$ and $\beta>0$.

LEMMA 2.3. *For each fixed $\beta\,(0<\beta\leqq 1)$ and each $x>0$ $(x\neq j_{\nu k},\ k=1,2,\cdots)$ the function $J_{\nu+\beta}(x)/J_\nu(x)$ decreases as ν increases, $-(\beta+1)/2\leqq\nu<\infty$, $\nu>-1$.*

Proof. Using (2.2), we find for $2\nu+\varepsilon+\beta>-1$,

$$\begin{aligned} &J_\nu(x)J_{\nu+\beta+\varepsilon}(x)-J_{\nu+\beta}(x)J_{\nu+\varepsilon}(x)\\ &\qquad=-\frac{4}{\pi}\int_0^{\pi/2}J_{2\nu+\varepsilon+\beta}(2x\cos\theta)\sin(\beta\theta)\sin(\varepsilon\theta)\,d\theta\\ &\qquad=-\left(\frac{2}{\pi x}\right)\int_0^{2x}J_{2\nu+\varepsilon+\beta}(u)\phi(u)\,du, \end{aligned} \tag{2.4}$$

where

$$u=2x\cos\theta\quad\text{and}\quad\phi(u)=\sin(\beta\theta)\sin(\varepsilon\theta)/\sin\theta.$$

As u increases from 0 to $2x$, θ decreases from $\pi/2$ to 0. For $0<\beta\leqq 1$ and $0<\varepsilon<1$, both $\sin(\beta\theta)$ and $\sin(\varepsilon\theta)/\sin\theta$ are increasing functions of θ on the interval $(0,\pi/2)$. Hence $\phi(u)$ decreases on $0<u<2x$ and by Lemma 2.1 the right hand side and hence the left hand side of (2.4) are negative. Now, if x is not a zero of $J_\nu(x)$ we can choose ε so small that $J_\nu(x)$ and $J_{\nu+\varepsilon}(x)$ have the same sign. Thus we get

$$J_{\nu+\beta+\varepsilon}(x)/J_{\nu+\varepsilon}(x)<J_{\nu+\beta}(x)/J_\nu(x)$$

and the lemma is proved.

LEMMA 2.4. *For each fixed $\beta>0$ and each fixed ν satisfying $\nu>-\beta/2$ the function*

$$K_{\nu+\beta}(x)/K_\nu(x)$$

decreases to 1 as x increases from 0 to ∞.

Proof. Here we use the recurrence relations [13, p. 79]

$$K_{\nu-1}(x)+K_{\nu+1}(x)=-2K'_\nu(x) \tag{2.5}$$

and

$$[K_{\nu-1}(x)-K_{\nu+1}(x)]=-2\nu K_\nu(x)/x. \tag{2.6}$$

From (2.1) and (2.5) we get

$$\begin{aligned} &K_\nu^2(x)[K_{\nu+\beta}(x)/K_\nu(x)]' \\ &\quad=2\int_0^\infty [K_{2\nu+\beta-1}(2x\cosh t)-K_{2\nu+\beta+1}(2x\cosh t)]\sinh(\beta t)\sinh t\,dt \\ &\quad=-2(2\nu+\beta)x^{-1}\int_0^\infty K_{2\nu+\beta}(2x\cosh t)\sinh(\beta t)\tanh t\,dt \end{aligned}$$

on using (2.6). The result follows easily by use of the positivity of K_ν for positive argument and the asymptotic formula for $K_\nu(x)$ [13, p. 202].

Using Sturm comparison techniques, L. Lorch [8] obtained this result as a special case of results on Whittaker functions but stated it only for the case $\nu \geqq 0$. However, as he has pointed out to us, his proof actually covers the full range of ν considered here. In the case $\nu \geqq 0$, this result was also given by P. Hartman and G. S. Watson as part of [5, Proposition 7.1].

LEMMA 2.5. *For each fixed β $(0<\beta\leqq 1)$ and each fixed ν satisfying $\nu \geqq -\beta/2$, the function*

$$J_{\nu+\beta}(x)/J_\nu(x)$$

increases with x in each interval $j_{\nu,k}<x<j_{\nu,k+1}$, $k=0,1,\cdots$.

Proof. The proof follows the same lines as that of Lemma 2.4. In case $\nu>-\beta/2$, the recurrence relation

$$2J_\nu'(x)=J_{\nu-1}(x)-J_{\nu+1}(x) \tag{2.7}$$

and (2.2) give

$$\begin{aligned}\pi J_\nu^2(x)[J_{\nu+\beta}(x)/J_\nu(x)]'=\int_0^{\pi/2} &[J_{2\nu+\beta-1}(2x\cos\theta) \\ &+J_{2\nu+\beta+1}(2x\cos\theta)]2\sin(\beta\theta)\sin\theta\,d\theta\end{aligned}$$

and this gives

$$\pi J_\nu^2(x)[J_{\nu+\beta}(x)/J_\nu(x)]'=\frac{2(2\nu+\beta)}{x}\int_0^{\pi/2} J_{2\nu+\beta}(2x\cos\theta)\sin(\beta\theta)\tan\theta\,d\theta$$

by use of the recurrence relation

$$2\nu J_\nu(x)=x[J_{\nu-1}(x)+J_{\nu+1}(x)].$$

Now if we put $u=2x\cos\theta$ we find that

$$\pi x^2 J_\nu^2(x)\{J_{\nu+\beta}(x)/J_\nu(x)\}'=(2\nu+\beta)\int_0^{2x} J_{2\nu+\beta}(u)\psi(u)\,du, \tag{2.8}$$

where $\psi(u)=\sin(\beta\theta)/\cos\theta$. But $\psi(u)$ is a decreasing function of u. Hence, by Lemma 2.1, the right hand side of (2.8) is positive and the result follows.

In case $\nu=-\beta/2$ we use

$$[J_{-\nu}(x)/J_\nu(x)]'=-2\sin(\nu\pi)/[\pi x J_\nu^2(x)]$$

where we have used a formula [13, p. 43] for the Wronskian of J_ν and $J_{-\nu}$. But this last expression is positive since $-\frac{1}{2} \leqq \nu < 0$.

We conclude this section with a variant of Lemma 2.4 which arises when we consider a ratio of modified Bessel functions multiplied by a suitable power of x.

LEMMA 2.6. *For each real ν and for each $\beta > 0$ the function*

$$x^\beta K_{\nu+\beta}(x)/K_\nu(x)$$

increases with x, $0 < x < \infty$.

Proof. We have

$$\begin{aligned}(d/dx)&[x^\beta K_{\nu+\beta}(x)/K_\nu(x)]\\ &= (d/dx)[x^{\nu+\beta} K_{\nu+\beta}(x)/\{x^\nu K_\nu(x)\}]\\ &= x^\beta [K_\nu(x)]^{-2} [K_{\nu+\beta}(x) K_{\nu-1}(x) - K_{\nu+\beta-1}(x) K_\nu(x)]\end{aligned} \tag{2.9}$$

on using [13, p. 79]

$$(d/dx)[x^\nu K_\nu(x)] = -x^\nu K_{\nu-1}(x).$$

Now using (2.1) we find that the right hand side of (2.9) is positive. Hence the result follows.

Remark. The results pertaining to K_ν (Lemmas 2.2, 2.4 and 2.6) hold for $\beta > 0$. Those for J_ν (Lemmas 2.3 and 2.5) have been proved only in the case $0 < \beta \leqq 1$ though it seems likely that they are valid for a somewhat larger range of values of β.

3. Zeros of a cross-product of Bessel functions. Here we describe the location of the positive zeros of (1.6) in the case $0 < \beta \leqq 1$ and their variation with ν and α.

THEOREM 3.1. *For fixed β ($0 < \beta \leqq 1$) fixed ν ($\geqq -\beta/2$) and fixed α (>0), the equation*

$$J_{\nu+\beta}(x)/J_\nu(x) = \alpha^\beta K_{\nu+\beta}(\alpha x)/K_\nu(\alpha x) \tag{3.1}$$

has infinitely many positive roots which we denote in increasing order by $\lambda(k, \nu, \alpha, \beta)$, $k = 1, 2, \cdots$. We have

$$j_{\nu+\beta,k-1} < \lambda(k, \nu, \alpha, \beta) < j_{\nu,k}, \qquad k = 1, 2, \cdots. \tag{3.2}$$

Moreover, for fixed ν, β and k, $\lambda(k, \nu, \alpha, \beta)$ increases from $j_{\nu+\beta,k-1}$ to $j_{\nu,k}$ as α increases from 0 to ∞.

Proof. The zeros of $J_\nu(x)$ and $J_{\nu+\beta}(x)$ interlace according to the pattern

$$0 < j_{\nu,1} < j_{\nu+\beta,1} < j_{\nu,2} < j_{\nu+\beta,2} < \cdots.$$

This follows from the fact that $j_{\nu k}$ increases with ν ($\nu > -1$) [13, p. 508] and

$$j_{\nu+\beta,k} \leqq j_{\nu+1,k} < j_{\nu,k+1}$$

[13, p. 479]. By Lemma 2.5 the left hand side of (3.1) increases between 0 and $j_{\nu 1}$ and between the positive zeros of $J_\nu(x)$ where it becomes infinite. In fact in each interval $(j_{\nu k}, j_{\nu+\beta,k})$, $k = 1, 2, \cdots$, the left hand side of (3.1) increases from $-\infty$ to 0 while in each interval $(j_{\nu+\beta,k}, j_{\nu,k+1})$, $k = 0, 1, \cdots$, it increases from 0 to ∞. By Lemma 2.4, the right hand side of (3.1) decreases as x increases on $(0, \infty)$ in case $\nu > -\beta/2$ and it is constant in case $\nu = -\beta/2$. This proves the first assertion of the theorem and the inequalities (3.2).

To see that the zeros of (3.1) increase with α, we simply observe that for each fixed x the right hand side of (3.1) increases with α (Lemma 2.6) while the left hand side does not change. When $\alpha \to 0+$ the zeros reduce to those of $J_{\nu+\beta}(x)$ while as $\alpha \to \infty$ they approach the zeros of $J_\nu(x)$.

Our principal result is a generalization of the Askey conjecture in the case $0<\beta \leqq 1$.

THEOREM 3.2. *With the notation of Theorem* 3.1, *for fixed* β $(0<\beta\leqq 1)$ *fixed* k $(=1, 2, \cdots)$ *and fixed* α (>0) *the root* $\lambda(k, \nu, \alpha, \beta)$ *increases with* ν, $-\beta/2 \leqq \nu < \infty$, $\nu > -1$.

Proof. Let ε be a small positive number and write λ_ν for $\lambda(k, \nu, \alpha, \beta)$. If we write

$$f(\nu, x) = \alpha^\beta \frac{K_{\nu+\beta}(\alpha x)}{K_\nu(\alpha x)} - \frac{J_{\nu+\beta}(x)}{J_\nu(x)}.$$

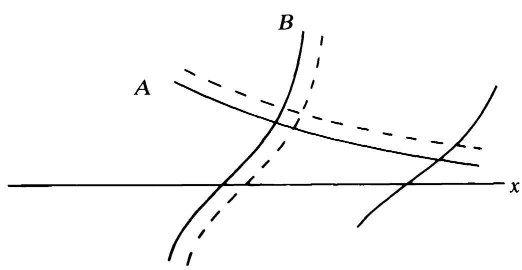

FIG. 1

we have, from Lemmas 2.2 and 2.3, that $f(\nu, x)$ increases with ν for fixed x. Hence

$$f(\nu+\varepsilon, \lambda_\nu) > f(\nu, \lambda_\nu) = 0.$$

But from Lemmas 2.4 and 2.5, $f(\nu+\varepsilon, x)$ is a decreasing function of x. Hence $\lambda_{\nu+\varepsilon} > \lambda_\nu$ and the theorem is proved.

The idea of the proof can be seen from the diagram (Fig. 1). The solid curves A and B represent, respectively, a part of the graph of the right hand side of (3.1) and a part of one branch of the graph of its left hand side. The broken curves represent parts of the same graphs for slightly larger values of ν. Lemmas 2.2 to 2.5 give the relative positions of these curves. Equation (3.1) holds and the cross-product (1.6) is zero for the value of x where the solid curves intersect. Clearly this value increases as ν increases.

We observe that the above argument can be used to prove the following more general result.

THEOREM 3.3. *Let* $\phi(x)$ *be a strictly increasing differentiable function on* $(0, \infty)$ *and* C *a positive constant. Then for fixed* β, $0<\beta\leqq 1$, *the positive zeros of fixed rank of*

$$J_{\nu+\beta}(x)/J_\nu(x) - CK_{\nu+\beta}[\phi(x)]/K_\nu[\phi(x)]$$

increase with ν, $-\beta/2 \leqq \nu < \infty$, $\nu > -1$.

4. The eigenvalue problem. Subject to the stated boundary conditions, the differential equation in (1.4) has solution

$$y(x) = \begin{cases} AJ_\nu(\lambda x), & 0 \leqq x \leqq 1, \\ BK_\nu(\alpha\lambda x), & x > 1. \end{cases}$$

If $y(x)$ and $y'(x)$ are to be continuous at $x=1$ we must have

$$AJ_\nu(\lambda)=BK_\nu(\alpha\lambda) \tag{4.1}$$

and

$$A\lambda J_\nu'(\lambda)=B\alpha\lambda K_\nu'(\alpha\lambda). \tag{4.2}$$

In view of the recurrence relations [13, pp. 18 and 79] and (4.1), (4.2) can be written

$$AJ_{\nu+1}(\lambda)=B\alpha K_{\nu+1}(\alpha\lambda). \tag{4.3}$$

There will be a nontrivial solution of (4.1), (4.3) for A and B if and only if

$$\alpha J_\nu(\lambda)K_{\nu+1}(\alpha\lambda)=J_{\nu+1}(\lambda)K_\nu(\alpha\lambda)$$

and in this case we may take

$$A=K_\nu(\alpha\lambda),\qquad B=J_\nu(\lambda).$$

Thus the boundary-value problem (1.4) has for its eigenvalues the zeros of the cross-product (1.5) and corresponding eigenfunctions

$$\psi_\nu(x)=\begin{cases}K_\nu(\alpha\lambda)J_\nu(\lambda x), & 0<x\leqq 1,\\ J_\nu(\lambda)K_\nu(\alpha\lambda x), & x\geqq 1.\end{cases} \tag{4.4}$$

Thus it is clear that Nagylaki's conjecture (1.3) follows from the case $\beta=1$ of Theorem 3.2.

The principal result of this section depends on a formula, (4.7) here, for the derivative of an eigenvalue with respect to a parameter. This formula is obtained by Sturmian-type arguments; it can be regarded as special case of the Hellmann–Feynman theorem of quantum chemistry; see [7] for further references and other applications of this idea to problems connected with zeros of Bessel functions.

The first part of the following theorem is a special case of Theorem 3.2 but is proved again since the approach in this section is quite different from that in § 3. The second part of the Theorem does not appear to be derivable by the methods used in § 3.

THEOREM 4.1. *With the notation of Theorem* 3.1, *for fixed* α (>0) *and fixed* k $(=1,2,\cdots)$, $\lambda(k,\nu,\alpha,1)$ *increases with* ν $(0\leqq\nu<\infty)$ *and* $\lambda(k,\nu,\alpha,1)/\nu$ *decreases to* 1 *as* ν *increases from* 0 *to* ∞.

Proof. If $\psi_\nu(x)$ is given by (4.4), we have

$$-(x\psi_\nu')'\psi_\nu+(\nu^2/x)\psi_\nu=\lambda^2gx\psi_\nu^2.$$

Integrating from 0 to ∞ we get, for $\nu>0$,

$$\lambda^2\int_0^\infty g(x)x\psi_\nu^2(x)\,dx=\nu^2\int_0^\infty x^{-1}\psi_\nu^2(x)\,dx+\int_0^\infty x[\psi_\nu'(x)]^2\,dx, \tag{4.5}$$

where integration by parts has been used in the last integral. This shows that the integral on the left hand side of (4.5) is positive. Thus we may normalize by choosing $\phi_\nu(x)=c_\nu\psi_\nu(\lambda)$ and

$$\int_0^\infty g(x)x\phi_\nu^2(x)\,dx=1. \tag{4.6}$$

We write λ_ν for $\lambda(k,\nu,\alpha,1)$; we multiply the equations

$$-(x\phi_\nu')'+(\nu^2/x)\phi_\nu=\lambda_\nu^2gx\phi_\nu$$
$$-(x\phi_\mu')'+(\mu^2/x)\phi_\mu=\lambda_\mu^2gx\phi_\mu$$

by ϕ_μ, ϕ_ν respectively, subtract and integrate between 0 and ∞ to get

$$x(\phi_\nu\phi_\mu' - \phi_\mu\phi_\nu')\Big|_0^\infty + (\nu^2-\mu^2)\int_0^\infty \frac{1}{x}\phi_\mu(x)\phi_\nu(x)\,dx$$
$$= (\lambda_\nu^2 - \lambda_\mu^2)\int_0^\infty g(x)x\phi_\mu(x)\phi_\nu(x)\,dx.$$

The integrated terms vanish at 0 and ∞ if $\mu > 0$, $\nu > 0$. Dividing by $\nu-\mu$ and letting $\mu \to \nu$ we get

(4.7)
$$\frac{d\lambda_\nu^2}{d\nu} = 2\nu\int_0^\infty \frac{1}{x}\phi_\nu^2(x)\,dx > 0.$$

This shows that λ_ν increases with ν for $\nu > 0$. But λ_ν is a continuous function of ν ($\nu \geqq 0$) since the terms in the cross-product (1.5) depend continuously on v for fixed x and α. Hence $\lambda(k, \nu, \alpha, 1)$ increases with ν for $\nu \geqq 0$.

We also have, for $\nu > 0$,

$$\frac{d}{d\nu}(\lambda_\nu^2/\nu^2) = \frac{1}{\nu^2}\frac{d}{d\nu}(\lambda_\nu^2) - \frac{2}{\nu^3}\lambda_\nu^2$$
$$= -2\nu^{-3}\int_0^\infty x[\phi_\nu'(x)]^2\,dx < 0$$

from (4.7), (4.5) and (4.6). Hence λ_ν/ν decreases with ν for $\nu > 0$. The fact that $\lambda_\nu/\nu \to 1$ as $\nu \to \infty$ follows from Theorem 3.1 and the known asymptotic formula for $j_{\nu,k}$ [12, p. 18].

It is clear that the methods of this section may be applied to boundary-value problems of the type (1.4) in which $g(x)$ need not have the special form given by (1.2).

Acknowledgment. The authors are grateful to Professor Lee Lorch for his interest in this work and for several helpful comments.

Note added in proof. A proof of the conjecture mentioned in the last paragraph of the Introduction is included in a paper by the first author, *Integral representations and complete monotonicity of various quotients of Bessel functions*, which is to appear in Canad. J. Math., December 1977.

REFERENCES

[1] R. G. COOKE. *Gibbs's phenomenon in Fourier–Bessel series and integrals*, Proc. London Math. Soc., 27 (1927), no. 2, pp. 171–192.

[2] ———, *A monotonicity property of Bessel functions*, J. London Math. Soc., 12 (1937), pp. 180–185.

[3] A. ERDÉLYI ET AL., *Higher Transcendental Functions*, vol. 2, McGraw-Hill, New York, 1955.

[4] E. GROSSWALD, *The Student t-distribution of any degree of freedom is infinitely divisible*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 36 (1976), pp. 103–109.

[5] P. HARTMAN AND G. S. WATSON, *"Normal" distribution functions on spheres and the modified Bessel functions*, Ann. Probability, 2 (1974), pp. 593–607.

[6] M. E. H. ISMAIL AND D. KELKER, *The Bessel polynomials and the Student t-distribution*, this Journal, 7 (1976), pp. 82–91.

[7] J. T. LEWIS AND M. E. MULDOON, *Monotonicity and convexity properties of zeros of Bessel functions*, this Journal, 8 (1977), pp. 171–178.

[8] L. LORCH, *Inequalities for some Whittaker functions*, Arch. Math. (Brno), 3 (1967), pp. 1–10.

[9] T. NAGYLAKI, *Conditions for the existence of clines*, Genetics, 80 (1975), pp. 595–615.

[10] J. STEINIG, *On a monotonicity property of Bessel functions*, Math. Z., 122 (1971), pp. 363–365.

[11] G. SZEGÖ, *On an inequality of Turán concerning Legendre polynomials*, Bull. Amer. Math. Soc., 54 (1948), pp. 401–405.

[12] F. TRICOMI, *Sulle funzioni di Bessel di ordine e argomento pressochè uguali*, Atti Accad. Sci. Torino Cl. Sci. Fis. Mat. Natur., 83 (1949), pp. 3–20.

[13] G. N. WATSON, *A Treatise on the Theory of Bessel Functions*, 2nd ed., Cambridge University Press, Cambridge, England, 1944.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978
© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0019 $01.00/0

BOUNDING EIGENVALUES OF ELLIPTIC OPERATORS*

J. R. KUTTLER† AND V. G. SIGILLITO†

Abstract. We present a method which bounds the error between an estimate for a point in the spectrum and an eigenvalue of self-adjoint elliptic operators. The method makes use of trial functions which need not satisfy the boundary conditions of the problem. By using rough preliminary estimates for the eigenvalues this method gives both upper and lower bounds.

1. Bounds for eigenvalues. In this paper we present a method which can be used to give upper and lower bounds for eigenvalues of self-adjoint elliptic operators including those of classical membrane and plate problems as well as Stekloff problems.

Our approach generalizes the inequalities of Fox, Henrici and Moler [3], Moler and Payne [9], and Nickel [10]. We prove a posteriori inequalities in a general Hilbert space setting which allows them to be combined with specific known a priori inequalities. The combined a posteriori-a priori inequality permits the estimation of eigenvalues in terms of quadratic forms of test functions. The test functions can be completely arbitrary except for smoothness conditions. In particular, no boundary conditions need be satisfied.

The a posteriori inequalities are given in the following theorem.

THEOREM. *Suppose A is an operator with domain $D(A)$ which is dense in the separable Hilbert space H. Let A be symmetric, so that*

$$(u, Av) = (Au, v) \qquad u, v \in D(A),$$

and let A have pure point spectrum $\{\lambda_i\}$ with corresponding orthonormal eigenvectors $\{u_i\}$ which are complete in H. Let A_* be an extension of A, so $D(A) \subset D(A_*) \subset H$ with $A_*u = Au$ for $u \in D(A)$.

For any number λ_*, and any $u_* \in D(A_*)$, suppose there exist w and h satisfying

$$A_*w = A_*u_* - \lambda_*u_*, \qquad w - u_* \in D(A), \tag{1}$$

$$A_*h = 0, \qquad h - u_* \in D(A). \tag{2}$$

Then

$$\min_i \left|\frac{\lambda_i - \lambda_*}{\lambda_i}\right| \leqq \frac{\|w\|}{\|u_*\|}, \tag{3}$$

$$\min_i \left|\frac{\lambda_i - \lambda_*}{\lambda_i}\right| \leqq \frac{\|\lambda_*h + A_*u_* - \lambda_*u_*\|}{\|A_*u_*\|}. \tag{4}$$

Proof. Since the eigenfunctions $\{u_i\}$ form a complete orthonormal set in H, by the Parseval identity

$$\|u_*\|^2 = \sum_i |(u_*, u_i)|^2,$$

* Received by the editors June 30, 1975, and in final revised form January 17, 1977.

† Applied Physics Laboratory, Johns Hopkins University, Laurel, Maryland 20810. This work was supported by the Department of the Navy, Naval Sea Systems Command under Contract N00017-72-C-4401.

and so,

$$\min_i \left|\frac{\lambda_* - \lambda_i}{\lambda_i}\right|^2 \|u_*\|^2 \leqq \sum_i \left|\frac{(\lambda_* - \lambda_i)(u_*, u_i)}{\lambda_i}\right|^2$$

$$= \sum_i \left|\frac{(A_*(u_* - w), u_i) - (u_*, Au_i)}{\lambda_i}\right|^2$$

$$= \sum_i \left|\frac{(u_* - w, Au_i) - (u_*, Au_i)}{\lambda_i}\right|^2$$

$$= \sum_i |(w, u_i)|^2 = \|w\|^2.$$

This proves (3). similarly,

$$\|A_* u_*\|^2 = \sum_i |(A_* u_*, u_i)|^2,$$

and so,

$$\min_i \left|\frac{\lambda_* - \lambda_i}{\lambda_i}\right|^2 \|A_* u_*\|^2 \leqq \sum_i \left|\frac{(\lambda_* - \lambda_i)(A_* u_*, u_i)}{\lambda_i}\right|^2$$

$$= \sum_i |(\lambda_* h + A_* u_* - \lambda_* u_*, u_i)|^2$$

$$= \|\lambda_* h + A_* u_* - \lambda_* u_*\|^2,$$

which proves (4).

It is not desirable to have to actually obtain the functions w and h which appear in the theorem and it is at this point that we introduce appropriate a priori inequalities which estimate w and h in terms of u_*.

2. An example. Let R be a bounded region of Euclidean n-space with boundary ∂R. Let the Hilbert space H be $\mathfrak{L}_2(R)$, the space of functions which are square integrable on R. Let A be the negative Laplacian $-\Delta$ with domain $D(A)$ the twice differentiable functions vanishing on ∂R. Then A_* will be $-\Delta$ with domain $D(A_*)$ functions which are just twice differentiable on R.

We have the classical fixed membrane problem

$$-\Delta u = \lambda u \quad \text{on } R, \qquad u = 0 \quad \text{on } \partial R, \tag{5}$$

whose eigenvalues and eigenvectors satisfy the hypotheses of the theorem. The appropriate a priori inequality to use is

$$\left(\int_R w^2\,dx\right)^{1/2} \leqq C_1\left(\int_R (\Delta w)^2\,dx\right)^{1/2} + C_2\left(\oint_{\partial R} w^2\,ds\right)^{1/2} \tag{6}$$

for $w \in D(A_*)$. See, e.g., [6, p. 153].

Put (6) into (3) and use (1) to obtain

$$\min_i \left|\frac{\lambda_* - \lambda_i}{\lambda_i}\right| \leqq \frac{C_1(\int_R (\Delta u_* + \lambda_* u_*)^2\,dx)^{1/2} + C_2(\int_{\partial R} u_*^2\,ds)^{1/2}}{(\int_R u_*^2\,dx)^{1/2}}$$

which yields

$$\min_i \left|\frac{\lambda_* - \lambda_i}{\lambda_i}\right|^2 \leqq \frac{2C_1^2 \int_R (\Delta u_* + \lambda_* u_*)^2\,dx + 2C_2^2 \oint_{\partial R} u_*^2\,ds}{\int_R u_*^2\,ds}. \tag{7}$$

Now the right side of (7) is a ratio of quadratic forms in the arbitrary twice-differentiable function u_*. Thus we can let u_* be a linear combination of test functions, say

$$u_* = \sum_{k=1}^{n} a_k \varphi_k, \tag{8}$$

and minimize the right side of (7) with respect to the coefficients a_k as in the Rayleigh–Ritz method. This leads to the relative matrix eigenvalue problem

$$Ma - \varepsilon^2 Na = 0, \tag{9}$$

where

$$M = \left[2C_1^2 \int_R (\Delta\varphi_i + \lambda_*\varphi_i)(\Delta\varphi_j + \lambda_*\varphi_j)\,dx + 2C_2^2 \oint_{\partial R} \varphi_i\varphi_j\,ds\right],$$

$$N = \left[\int_R \varphi_i\varphi_j\,dx\right],$$

$$a = (a_1, \cdots, a_n)^T.$$

Now let ε be the smallest eigenvalue of (9); then

$$\min_i \left|\frac{\lambda_* - \lambda_i}{\lambda_i}\right| \leqq \varepsilon,$$

or, if $\varepsilon < 1$,

$$\lambda_*/(1+\varepsilon) \leqq \lambda_i \leqq \lambda_*/(1-\varepsilon),$$

which gives upper and lower bounds for the eigenvalue λ_i which is closest to λ_*.

Similarly, by using the triangle inequality on (4), employing (6), and then using the arithmetic-geometric mean inequality, we obtain

$$\min_i \left|\frac{\lambda_* - \lambda_i}{\lambda_i}\right|^2 \leqq \frac{2\int_R (\Delta u_* + \lambda_* u_*)^2\,dx + 2C_2^2\lambda_*^2 \oint_{\partial R} u_*^2\,ds}{\int_R (\Delta u_*)^2\,dx}, \tag{10}$$

which may be used as an alternate to (7) with the above technique.

3. Application. To use the method of the previous section requires several things. First, explicit values for the constants C_1, C_2 in (6) are needed. The optimum values for these constants are [6]

$$C_1 = \lambda_1^{-1}, \qquad C_2 = q_1^{-1/2}.$$

Here q_1 is the lowest eigenvalue of the Dirichlet eigenvalue problem [12]

$$\begin{aligned} &\Delta^2 v = 0 \quad \text{in } R, \\ &v = \Delta v - q\frac{\partial v}{\partial n} = 0 \quad \text{on } \partial R, \end{aligned} \tag{11}$$

where $\partial/\partial n$ is the normal derivative. Exact values of λ_1, q_1 are not needed, only lower bounds. In this connection see [4], [5], [6], [11], [12].

Next, a method of guessing λ_* is required. The inequality (10) says only that some eigenvalue λ_i is in the interval $(\lambda_*/(1+\varepsilon), \lambda_*(1-\varepsilon))$. If, say, λ_1 is the desired eigenvalue, then a crude lower bound on λ_2 is required to start the procedure. Improved guesses for λ_* are generated as follows. Let u_* be given by (8) where the a_i satisfy (9). Holding u_* fixed, minimize (7) with respect to λ_*. This results in

$$\lambda_* = -\frac{\sum_{i,j=1}^{n} a_i a_j \int_R \varphi_i \, \Delta\varphi_j \, dx}{\sum_{i,j=1}^{n} a_i a_j \int_R \varphi_i \varphi_j \, dx}. \tag{12}$$

If the value in (12) is less than the lower bound on λ_2, it can be taken as the new guess for λ_*, and the procedure can then be repeated.

4. Numerical results. The method was numerically tested on rhombical membranes. These were chosen because excellent tables of upper and lower bounds for selected eigenvalues have been computed by Stadter [15]. The Weinstein–Aronszajn method of intermediate problems with the *special choice* of Bazley and Fox was used with carefully constructed trial functions.

Fox, Henrici and Moler, using Bessel functions in the method described in [3], also treated rhombical membranes, but were "unable to obtain bounds significantly tighter than those obtained in [15]."

We have obtained bounds comparable to Stadter's using only polynomials or elementary trigonometric functions.

Consider the rhombus with unit side and least interior angle θ as shown in Fig. 1. The test functions used were $x^{2i}y^{2j}$, $i, j = 0, 1, 2, \cdots, 8$, for $\theta = 75°$ and $\theta = 45°$, and $\cos(m\pi x)\cos(n\pi y)$, $m, n = 0, 1, 2, \cdots, 6$ for $\theta = 15°$. Notice that the trial functions

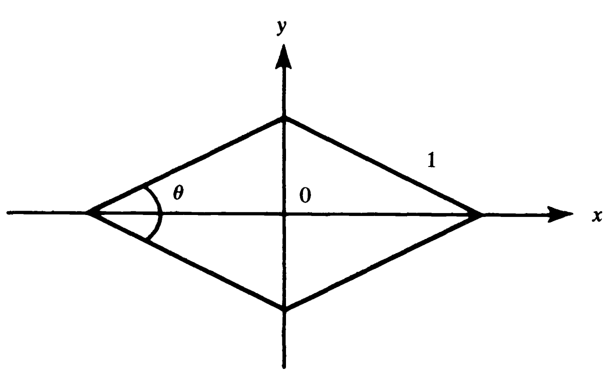

FIG. 1.

do not have to satisfy the boundary conditions. The results for the first three eigenvalues in the even-even symmetry class are given in Table 1. For comparison, Stadter's results are given in Table 2. (Note: our angle θ is the complement of Stadter's and his eigenvalues in [15] are divided by π^2.) The inequality (10) was used with angles 75° and 45°, but inquality (7) gave better results with 15°.

5. Other examples.

I. The free membrane eigenvalue problem is

$$\begin{aligned} \Delta u + \mu u &= 0 \quad \text{in } R, \\ \frac{\partial u}{\partial n} &= 0 \quad \text{on } \partial R, \end{aligned} \tag{13}$$

TABLE 1
Bounds for eigenvalues of unit side rhombical membrane in even-even symmetry class.

θ	75°		45°		15°	
i	lower	upper	lower	upper	lower	upper
1	20.8660	20.8782	34.7017	34.8569	196.8111	206.7539
2	79.0234	79.0711	100.1116	100.4642	340.7600	386.9026
3	108.8175	108.9697	183.3045	187.6383	436.9660	652.2823

TABLE 2
Stadter's bounds for eigenvalues of unit side rhombical membrane in even-even symmetry class.

θ	75°		45°		15°	
i	lower	upper	lower	upper	lower	upper
1	20.8613	20.8871	34.7113	35.0283	199.0205	207.6664
2	78.9173	79.2392	100.1073	101.2326	357.7928	390.1554
3	108.7433	109.2763	184.1569	189.7827	371.5017	708.1836

with eigenvalues $0=\mu_1<\mu_2\leqq\mu_3\leqq\cdots$. Combining inequality (3) with the appropriate a priori inequality yields

$$\min_{i\neq 1}\left|\frac{\mu_*-\mu_i}{\mu_i}\right|\leqq\frac{C_1(\int_R(\Delta u_*+\mu_*u_*)^2\,dx)^{1/2}+C_2(\oint_{\partial R}(\partial u_*/\partial n)^2\,ds)^{1/2}}{(\int u^2\,dx)^{1/2}} \tag{14}$$

for any number μ_* and any function u_* satisfying

$$\int_R u_*\,dx=0.$$

In connection with the values of C_1 and C_2 see particularly [1] and also [6], [7], [8], [11].

II. The clamped plate eigenvalue problem is

$$\begin{aligned}\Delta^2u-\Omega u&=0\quad\text{in } R,\\ u=\frac{\partial u}{\partial n}&=0\quad\text{on }\partial R,\end{aligned} \tag{15}$$

with eigenvalues $0<\Omega_1\leqq\Omega_2\leqq\cdots$. Combining (3) with the a priori inequality of [13] yields

$$\begin{aligned}&\min_i\left|\frac{\Omega_*-\Omega_i}{\Omega_i}\right|^2\int_R u^2\,dx\\ &\leqq\alpha_1\int_R(\Delta^2u_*-\Omega_*u_*)^2\,dx+\alpha_2\oint_{\partial R}u_*^2\,ds+\alpha_3\oint_{\partial R}\left(\frac{\partial u_*}{\partial n}\right)^2ds\\ &\qquad+\alpha_4\oint_{\partial R}\left(\frac{\partial u_*}{\partial t}\right)^2ds,\end{aligned} \tag{16}$$

for any number Ω_* and any function u_*. Here $\partial/\partial t$ is the tangential derivative on ∂R. For values of the constants see also [14].

III. The Stekloff eigenvalue problem is

$$(17)\qquad \Delta u = 0 \quad \text{in } R, \qquad \frac{\partial u}{\partial n} = pu \quad \text{on } \partial R,$$

with eigenvalues $0 = p_1 < p_2 \leqq p_3 \leqq \cdots$. Combining (3) with the appropriate a priori inequality yields

$$(18)\qquad \min_{i \neq 1} \left| \frac{p_* - p_i}{p_i} \right| \leqq \frac{C_1(\oint_{\partial R} (\partial u_*/\partial n - p_* u_*)^2 \, ds)^{1/2} + C_2(\int_R (\Delta u_*)^2 \, dx)^{1/2}}{\oint_{\partial R} u_*^2 \, ds}$$

for any number p_* and any function u_* satisfying

$$\oint_{\partial R} u_* \, ds = 0.$$

In connection with the values of C_1 and C_2 see particularly [1] and also [6], [7], [8], [12].

Other examples can be given, but the above should indicate the variety of problems which can be treated by the method. Further numerical results will be reported in forthcoming papers.

Acknowledgment. We would like to thank a referee for suggesting inequality (10).

REFERENCES

[1] J. H. BRAMBLE AND L. E. PAYNE, *Bounds in the Neumann problem for second order uniformly elliptic operators*, Pacific J. Math., 12 (1962), pp. 823–833.

[2] D. W. FOX AND W. C. RHEINBOLDT, *Computational methods for determining lower bounds for eigenvalues of operators in Hilbert space*, SIAM Rev., 8 (1966), pp. 427–462.

[3] L. FOX, P. HENRICI AND C. MOLER, *Approximations and bounds for eigenvalues of elliptic operators*, SIAM J. Numer. Anal., 4 (1967), pp. 89–102.

[4] J. HERSCH AND L. E. PAYNE, *One-dimensional auxiliary problems and a priori bounds*, Abh. Math. Sem. Univ. Hamburg, 36 (1971), pp. 57–65.

[5] J. R. KUTTLER, *Remarks on a Stekloff eigenvalue problem*, SIAM J. Numer. Anal., 9 (1972), pp. 1–5.

[6] J. R. KUTTLER AND V. G. SIGILLITO, *Inequalities for membrane and Stekloff eigenvalues*, J. Math. Anal. Appl., 23 (1968), pp. 148–160.

[7] ———, *Lower bounds for Stekloff and free membrane eigenvalues*, SIAM Rev., 10 (1968), pp. 368–370.

[8] ———, *An inequality for a Stekloff eigenvalue by the method of defect*, Proc. Amer. Math. Soc., 20 (1969), pp. 357–360.

[9] C. B. MOLER AND L. E. PAYNE, *Bounds for eigenvalues and eigenvectors of symmetric operators*, SIAM J. Numer. Anal., 5 (1968), pp. 64–70.

[10] K. L. E. NICKEL, *Extension of a recent paper by Fox, Henrici, and Moler on eigenvalues of elliptic operators*, Ibid., 4 (1967), pp. 483–488.

[11] L. E. Payne, *Isoperimetric inequalities and their applications*, SIAM Rev., 9 (1967), pp. 453–488.

[12] ———, *Some isoperimetric inequalities for harmonic functions*, this Journal, 1 (1970), pp. 354–359.

[13] V. G. SIGILLITO, *A priori inequalities and pointwise bounds for solutions of fourth order elliptic partial differential equations*, SIAM J. Appl. Math., 15 (1967), pp. 1136–1155.

[14] ———, *A priori inequalities and approximate solutions of the first boundary value problem for* $\Delta^2 u = f$, SIAM J. Numer. Anal., 13 (1976), pp. 251–260.

[15] J. T. STADTER, *Bounds to eigenvalues of rhombical membranes*, SIAM J. Appl. Math., 14 (1966), pp. 324–341.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0020 $01.00/0

A LOWER BOUND FOR THE EIGENVALUES OF THE ELLIPTIC DIRICHLET PROBLEM FOR A GENERAL DOMAIN IN TERMS OF ITS CHARACTERISTIC DIMENSION*

SHLOMO BREUER† AND JOSEPH J. ROSEMAN†

Abstract. We consider the Dirichlet eigenvalue problem on a general bounded domain $\mathcal{D} \subset R^n$, with a sufficiently regular boundary, for the equation $Lv - \lambda rv = 0$, where L is a linear, uniformly elliptic, self-adjoint differential operator of order $2s$, and r is a positive piecewise continuous function on $\mathcal{D}$. Quantities h, p, and δ are defined, where h is the supremum of the radii of all spheres contained in $\mathcal{D}$, p is an integer which characterizes its geometry, and $\delta = p^{1/(2s)}h$; δ is called the characteristic dimension of $\mathcal{D}$ and is, in general, independent of the overall size of $\mathcal{D}$. It can then be shown that $\lambda_1 \geqq M(\kappa/K)\delta^{-2s}$, where λ_1 is the smallest eigenvalue, κ is a parameter depending upon L, K is an upper bound for r and M depends only upon s and n. A detailed proof is given for $L = -\nabla^2$ and $r = 1$ in two dimensions and the extension to the general case is straightforward.

1. Introduction. For a domain $\mathcal{D} \subset R^n$ with a sufficiently regular boundary, $\partial\mathcal{D}$, we consider the following elliptic eigenvalue problem:

$$Lv - \lambda r(x)v = 0, \quad x \in \mathcal{D}, \tag{1.1a}$$

(1.1b) $\qquad v$ satisfies homogeneous Dirichlet data on $\partial\mathcal{D}$,

where $x = (x_1, x_2, \cdots, x_n)$, and

(i) L is a linear, uniformly strongly elliptic, self-adjoint operator of order $2s$ such that if T is the subset of $C^{2s}(R^n)$ which contains all nontrivial functions which are of compact support in $\mathcal{D}$, there exists a $\kappa > 0$ for which the following is valid for all $\psi \in T$:

$$\int_{\mathcal{D}} \psi L\psi \, dx \geqq \kappa \int_{\mathcal{D}} \sum_{i_1+i_2+\cdots+i_n=s} \left| \frac{\partial^s \psi}{\partial x_1^{i_1} \partial x_2^{i_2} \cdots \partial x_n^{i_n}} \right|^2 dx, \tag{1.2}$$

(ii) r is piecewise continuous and positive in $\mathcal{D}$ and there exists $K > 0$ such that

$$0 < r(x) < K, \qquad x \in \mathcal{D}. \tag{1.3}$$

In general κ and K may depend upon the domain $\mathcal{D}$. However, there is a wide class of situations for which κ and K are independent of $\mathcal{D}$; *in particular, this is true if r and the coefficients of L are constants.*

Much attention has been directed to the question of obtaining a positive lower bound for the first (lowest) eigenvalue, λ_1, and to the dependence of λ_1 on the domain $\mathcal{D}$ (cf. [1]–[11]). From dimensional considerations, we see that a lower bound estimate is expected to be of the form

$$\lambda_1 \geqq \frac{\kappa}{K} \frac{M_1}{h^{2s}}, \tag{1.4}$$

where h would have the dimension of length and would depend upon the size of $\mathcal{D}$, while M_1 would be dimensionless and would depend upon the geometry of $\mathcal{D}$ and possibly upon L, but not upon the size of $\mathcal{D}$.

For example, for $n = 2$, if $\mathcal{D}$ is a rectangle with sides d and D, with $d \leqq D$, and $L = -\nabla^2$, it is well known that

$$\lambda_1 = \pi^2(d^{-2} + D^{-2}) > \pi^2 d^{-2}. \tag{1.5}$$

* Received by the editors August 3, 1976 and in revised form January 26, 1977.
† Department of Mathematical Sciences, Tel-Aviv University, Tel-Aviv, Israel.

Since h depends upon the size of $\mathscr{D}$, the question arises as to how to find the smallest possible value of h for which (1.4) holds, and to describe, as explicitly as possible, the dependence of M_1 on the geometry of $\mathscr{D}$.

Let us now consider a bounded domain $\mathscr{D}$ such that the infimum of the radii of all discs which contain $\mathscr{D}$ is H and such that the supremum of the radii of all discs which are contained in $\mathscr{D}$ is h, with $h \ll H$. It is clear that one can expect the estimate

$$\lambda_1 \geqq \frac{\kappa}{K} \frac{M_2}{H^{2s}} \tag{1.6}$$

to hold with M_2 independent of the size of $\mathscr{D}$. In fact, it is well known that for convex domains one can obtain a much stronger estimate of the form

$$\lambda_1 \geqq \frac{\kappa}{K} \frac{M_2}{h^{2s}}. \tag{1.7}$$

Now it is the purpose of this paper to characterize a very wide class of domains, including nonconvex and multiply connected domains, for which an estimate of the form

$$\lambda_1 \geqq \frac{\kappa}{K} \frac{M_3}{ph^{2s}} \tag{1.8}$$

is valid, where p is an integer which characterizes the geometry of $\mathscr{D}$ (as explained below), but is independent of its size, and M_3 depends only upon s and n.

Defining

$$\delta = p^{1/(2s)} h, \tag{1.9}$$

we may consider δ to be the "characteristic dimension" of the domain with respect to the operator L for the problem (1.1) in that it takes into account both the "thickness" and the complexity of the geometry of $\mathscr{D}$.

Working independently, W. K. Hayman [4] recently obtained results similar to ours, which are to appear elsewhere. A comparison of our results and techniques with Hayman's appears below in § 4.

2. Preliminaries. We consider a domain $\mathscr{D} \subset R^n$, with boundary $\partial\mathscr{D}$ and closure $\bar{\mathscr{D}}$. We suppose that $\partial\mathscr{D}$ is a disjoint union of a finite number of closed, piecewise smooth hypersurfaces, and denote by h the supremum of the radii of all discs which are contained in $\mathscr{D}$. Next we shall proceed to define the "covering number" of $\mathscr{D}$.

Through every point A in $\mathscr{D}$, draw a line segment $\mathscr{L}_A$ with endpoints P_1 and P_2 such that P_1 lies on $\partial\mathscr{D}$, P_2 is in $\mathscr{D}$, the length of $\overline{AP_1}$ is less than $2h$, and the length, $\beta_1 > 0$, of $\overline{AP_2}$ is less than h. Consider a local Cartesian coordinate system, $\xi = (\xi_1, \xi_2, \cdots, \xi_n)$ centered at A, with $\mathscr{L}_A$ lying along the ξ_1-axis. Construct a domain $\mathscr{B}_A$ of the form

$$\mathscr{B}_A = [(\xi_1, \xi_2, \cdots, \xi_n) | \, |\xi_i| < m_A \quad (i = 2, 3, \cdots, n); \; -\beta_1 < \xi_1 < g(\xi_2, \xi_3, \cdots, \xi_n) + \beta_2], \tag{2.1}$$

where the hypersurface $\xi_1 = g(\xi_2, \xi_3, \cdots, \xi_n)$ is part of the boundary $\partial\mathscr{D}, 0 < g(\xi_2, \xi_3, \cdots, \xi_n) < 5h/2$ for $|\xi_i| < m_A$, and β_2 is a positive number less than $\frac{1}{2}h$.

We see that $\bar{\mathscr{D}} \subset \bigcup_{A \in \mathscr{D}} \mathscr{B}_A$. Therefore, by the Heine–Borel theorem, we may select a finite set of points $A_1, A_2, A_3, \cdots, A_N$ such that

$$\bar{\mathscr{D}} \subset \bigcup_{k=1}^{N} \mathscr{B}_{A_k}. \tag{2.2}$$

Consequently, there exists an integer $p \leqq N$ such that every point in $\mathscr{D}$ belongs to no more than p of the domains $\mathscr{B}_{A_k}$.

The value of the covering number of $\mathscr{D}$ is defined to be the smallest possible integer p that can be obtained for any construction of the above type.

We finally define the characteristic dimension δ:

$$\delta = p^{1/(2s)}h. \tag{2.3}$$

3. Statement and proof of the main theorem. For simplicity, the theorem and proof that are given below are stated for $n=2$, $L=-\nabla^2$, and $r(x)=1$. The generalization to the problem (1.1)–(1.3) which leads to the result (1.8) is straightforward.

MAIN THEOREM. *Suppose $\mathscr{D}$ is a bounded domain in R^2 whose boundary, $\partial\mathscr{D}$, is a disjoint union of a finite number of closed, piecewise smooth curves. Let $\mathscr{D}$ have the covering number p and characteristic dimension δ. Then if λ_1 is the first eigenvalue for the Laplacian operator with homogeneous Dirichlet boundary data, we have*

$$\lambda_1 \geqq \frac{\pi^2}{64\delta^2}. \tag{3.1}$$

Proof. The eigenvalue λ_1 can be characterized by means of the Rayleigh quotient as

$$\lambda_1 = \inf_{\psi\in T} \frac{(-\nabla^2\psi, \psi)}{(\psi, \psi)} = \inf_{\psi\in T} \frac{(\nabla\psi, \nabla\psi)}{(\psi, \psi)}, \tag{3.2}$$

where $(w_1, w_2) = \int_{\mathscr{D}} w_1(x_1, x_2)w_2(x_1, x_2)\,dx_1\,dx_2$, and T is that subset of $C^2(R^2)$ consisting of all nontrivial functions which have compact support in $\mathscr{D}$.

We shall now proceed to obtain a lower bound for the quotient $(\nabla\psi, \nabla\psi)/(\psi, \psi)$ which is valid for all $\psi \in T$.

The definition of the covering number implies that there exist a finite number of domains, $\mathscr{B}_{A_1}, \mathscr{B}_{A_2}, \cdots, \mathscr{B}_{A_N}$, such that

(3.3) (i)
$$\bar{\mathscr{D}} \subset \bigcup_{k=1}^{N} \mathscr{B}_{A_k},$$

(ii) every point in $\mathscr{D}$ belongs to at most p of these domains, and

(iii) in each $\mathscr{B}_{A_k}$ it is possible to construct a local Cartesian coordinate system, ξ_1^k-ξ_2^k, centered at A_k, such that $\mathscr{B}_{A_k}$ is a domain of the form

$$\mathscr{B}_{A_k} = [(\xi_1^k, \xi_2^k) \mid |\xi_2^k| < m_{A_k};\ -\beta_1^k < \xi_1^k < g^k(\xi_2^k) + \beta_2^k], \tag{3.4}$$

with the boundary curve $\xi_1^k = g^k(\xi_2^k)$ belonging to $\partial\mathscr{D}$, $0 < g^k(\xi_2^k) < 5h/2$ when $|\xi_2^k| < m_{A_k}$, β_1^k and β_2^k being positive numbers such that $\beta_1^k < h$ and $\beta_2^k < \frac{1}{2}h$.

Since $\psi = 0$ when $\xi_1^k = g^k(\xi_2^k)$, and since $|g^k + \beta_2^k + \beta_1^k| < 4h$, it follows that

$$\int_{\mathscr{B}_{A_k}} |\psi|^2\,d\xi_1^k\,d\xi_2^k \leqq \frac{64h^2}{\pi^2} \int_{\mathscr{B}_{A_k}} \left|\frac{\partial\psi}{\partial\xi_1^k}\right|^2 d\xi_1^k\,d\xi_2^k \leqq \frac{64h^2}{\pi^2} \int_{\mathscr{B}_{A_k}} |\nabla\psi|^2\,d\xi_1^k\,d\xi_2^k. \tag{3.5}$$

(The standard inequality

$$\int_0^{\tau_1} F^2(\tau)\,d\tau \leqq \frac{4\tau_1^2}{\pi^2} \int_0^{\tau_1} F'^2(\tau)\,d\tau,$$

valid for any function F which is piecewise differentiable on $[0, \tau_1]$ and such that $F(\tau) = 0$ for at least one point on $[0, \tau_1]$, was used above.)

A transformation from the local ξ_1^k-ξ_2^k system back to the global x_1-x_2 system yields

$$\int_{\mathscr{B}_{A_k}} |\psi|^2 \, dx_1 \, dx_2 \leqq \frac{64h^2}{\pi^2} \int_{\mathscr{B}_{A_k}} |\nabla\psi|^2 \, dx_1 \, dx_2. \tag{3.6}$$

From (3.3), we have that for any square integrable function $w(x_1, x_2)$,

$$\int_{\mathscr{D}} w^2(x_1, x_2) \, dx_1 \, dx_2 \leqq \sum_{k=1}^{N} \int_{\mathscr{B}_{A_k}} w^2(x_1, x_2) \, dx_1 \, dx_2. \tag{3.7}$$

Suppose now that $w = 0$ in the exterior of $\mathscr{D}$. Then, since every point of $\mathscr{D}$ belongs to at most p sets of the family $\{\mathscr{B}_{A_k}\}$, we also have

$$\sum_{k=1}^{N} \int_{\mathscr{B}_{A_k}} w^2(x_1, x_2) \, dx_1 \, dx_2 \leqq p \int_{\mathscr{D}} w^2(x_1, x_2) \, dx_1 \, dx_2. \tag{3.8}$$

From (3.6), (3.7), and (3.8) it follows that

$$\int_{\mathscr{D}} |\psi|^2 \, dx_1 \, dx_2 \leqq \frac{64h^2p}{\pi^2} \int_{\mathscr{D}} |\nabla\psi|^2 \, dx_1 \, dx_2. \tag{3.9}$$

Thus,

$$\frac{(\nabla\psi, \nabla\psi)}{(\psi, \psi)} \geqq \frac{\pi^2}{64\delta^2}, \tag{3.10}$$

for all $\psi \in T$. Therefore,

$$\lambda_1 = \inf_{\psi \in T} \frac{(\nabla\psi, \nabla\psi)}{(\psi, \psi)} \geqq \frac{\pi^2}{64\delta^2}, \tag{3.11}$$

and the proof is complete.

4. Additional remarks. As mentioned earlier, Hayman [4] has recently obtained results similar to ours, also by means of a covering procedure. His technique is to cover the domain in question with circles whose centers lie on the boundary, in contrast to our method of covering with rectangle-like domains constructed about interior points.

Hayman showed that for two-dimensional simply connected domains, the first eigenvalue for the Laplacian operator satisfies

$$\lambda > \frac{1}{900h^2}. \tag{4.1}$$

Comparing (4.1) and (3.1), we see that in the context of two-dimensional simply connected domains, our results yield a larger bound when $p \leqq 138$, while Hayman's uniform bound is larger for $p > 138$.

Hayman also obtains a bound for the first eigenvalue of the Laplacian operator in n dimensions. However, here his bound is nonuniform in that the constant depends upon the geometry of the domains, as does our p.

Hayman's methods, as given in [4], do not apply to all multiply connected domains, even in two dimensions. A simple example is the annulus

$$\mathscr{G} = [(x, y) \,|\, a^2 < x^2 + y^2 < R^2],$$

with a/R sufficiently small.

Our methods, on the other hand, do apply to $\mathscr{G}$ and in fact to all multiply connected domains, whose boundary is sufficiently regular.

The work in this paper, as well as in [4], has been for the Dirichlet problem. M. Bareket has privately communicated to us that with the aid of our covering techniques, she has obtained lower estimates for the mixed boundary value problem for the Laplacian operator in two dimensions. Those results will appear elsewhere.

Acknowledgment. We express our thanks to the referees and editors for having brought W. K. Hayman's work to our attention, and for their constructive criticism.

REFERENCES

[1] M. BAREKET AND B. RULF, *An eigenvalue problem related to sound propagation in elastic tubes*, J. Sound and Vibration, 38(1975), pp. 437–449.

[2] A. BITSADZE, *Boundary Value Problems for Second Order Elliptic Equations*, North-Holland, Amsterdam, 1968.

[3] G. FICHERA, *Linear Elliptic Differential Systems and Eigenvalue Problems*, Springer-Verlag, Berlin, 1965.

[4] W. K. HAYMAN, *Some bounds for principal frequency*, Applicable Anal., to appear.

[5] C. O. HORGAN, *Inequalities of Korn and Friedrichs in elasticity and potential theory*, Z. Angew. Math. Phys., 26 (1975), pp. 155–164.

[6] C. O. HORGAN AND J. K. KNOWLES, *Eigenvalue problems associated with Korn's inequalities*, Arch. Rational Mech. Anal., 40 (1971), pp. 384–401.

[7] L. E. PAYNE AND H. F. WEINBERGER, *Lower bounds for vibration frequencies of elastically supported membranes and plates*, J. Soc. Indust. Appl. Math., 5 (1957), pp. 171–182.

[8] M. H. PROTTER, *Lower bounds for the first eigenvalue of elliptic equations*, Ann. of Math., 71 (1960), pp. 423–444.

[9] M. H. PROTTER AND H. F. WEINBERGER, *On the spectrum of general second order operators*, Bull. Amer. Math. Soc., 72 (1966), pp. 251–255.

[10] H. WEINBERGER, *Variational Methods for Eigenvalue Approximation*, Regional Conference Series in Applied Mathematics, no. 15, Society for Industrial and Applied Mathematics, Philadelphia, 1974.

[11] R. WEINSTOCK, *Inequalities for a classical eigenvalue problem*, J. Math. Mech., 3 (1954), pp. 745–753.

SIAM J. MATH. ANAL.
Vol. 9, No. 4, August 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0904-0021 $01.00/0

CORRIGENDUM: NEW RELATIONS BETWEEN TWO TYPES OF BESSEL FUNCTION INTEGRALS*

HENRY E. FETTIS†

The following acknowledgment should appear as a footnote: "This work was supported in part by Flight Dynamics Laboratory, U.S. Air Force, Wright–Patterson AFB, Ohio."

* This Journal, 8 (1977), pp. 978–982. Received by the editors December 19, 1977.
† 1885 California, Apartment 62, Mountain View, California 94041.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978
© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0001 $01.00/0

SOLUTION OF CERTAIN RECURRENCES. II*

L. CARLITZ†

Abstract. The function $\{n; \mathbf{r}, \mathbf{t}\}$, where $\mathbf{r}=(r_1, \cdots, r_k)$, $\mathbf{t}=(t_1, \cdots, t_k)$, $\boldsymbol{\delta}=(\delta_1, \cdots, \delta_k)$, is defined by means of $\{n; \mathbf{r}, \mathbf{t}\}=0$ if $r_1+\cdots+r_k>n$ or any $r_j<0$; $\{0; \mathbf{0}; \mathbf{t}\}=1$; $\{n+1; \mathbf{r}; \mathbf{t}\}=\sum\{n; \mathbf{r}-\boldsymbol{\delta}; \mathbf{t}\}$, where $\delta_j=0, 1, \cdots, t_j$. The function $\{n; \mathbf{r}; \mathbf{t}\}$ is evaluated in the present paper. The special case $t_1=\cdots=t_k$ had been considered earlier.

A combinatorial application of $\{n; \mathbf{r}, \mathbf{t}\}$ to the enumeration of rectangular matrices satisfying certain restrictions is given at the end of the paper.

1. Let $n, k, t, r_1, \cdots, r_k$ be integers, $n\geqq 0$, $k\geqq 1$, $t\geqq 1$. Define the function $\{n; r_1, \cdots, r_k\}_t$ by means of

$$\{n; r_1, \cdots, r_k\}_t=0 \quad \text{if } r_1+\cdots+r_k>n \tag{1.1}$$

or, if any $r_j<0$,

$$\{0; 0, \cdots, 0\}_t=1, \tag{1.2}$$

and

$$\{n+1; r_1, \cdots, r_k\}_t=\sum\{n; r_1-\delta_1, \cdots, r_k-\delta_k\}, \tag{1.3}$$

where each $\delta_j=0, 1, \cdots, t$; thus the right hand side of (1.3) contains $(t+1)^k$ summands. Generalizing some results of Narayana and Rohatgi [4], the writer [1] proved the following.

Put

$$(1+x+\cdots+x^t)^n=\sum_{k=0}^{\infty} c_t(n, k)x^k \tag{1.4}$$

and

$$\phi_t(x)(1+x+\cdots+x^t)^n=\sum_{k=0}^{\infty} f_t(n, k)x^k, \tag{1.5}$$

where

$$\phi_t(x)=\left(1-x\frac{d}{dx}\right)(1+x+\cdots+x^t). \tag{1.6}$$

Then, for $k=1$,

$$\{n; r\}_t=f_t(n, r) \quad \text{for } 0\leqq r\leqq n, \tag{1.7}$$

so that

$$\phi_t(x)(1+x+\cdots+x^t)^n=\sum_{r=0}^{n}\{n; r\}_t x^r+O(x^{n+2}), \tag{1.8}$$

where $O(x^{n+2})$ denotes a polynomial divisible by x^{n+2}.

For $k\geqq 1$ put

$$\prod_{i=1}^{k}(1+x_i+\cdots+x_i^t)^n=\sum_{r_1,\cdots,r_k=0}^{\infty} c_t(n; r_1, \cdots, r_k)x_1^{r_1}\cdots x_k^{r_k}, \tag{1.9}$$

* Received by the editors September 9, 1976, and in revised form December 31, 1976.
† Department of Mathematics, Duke University, Durham, North Carolina 27706.

so that, by (1.4),

$$c_t(n; r_1, \cdots, r_k) = \prod_{i=1}^{k} c_t(n, r_i). \tag{1.10}$$

Also define

$$\phi_t(x_1, \cdots, x_k) = \left(1 - x_1\frac{\partial}{\partial x_1} - \cdots - x_k\frac{\partial}{\partial x_k}\right) \prod_{i=1}^{k} (1 + x_i + \cdots + x_i^t) \tag{1.11}$$

and put

$$\phi_t(x_1, \cdots, x_k) \prod_{i=1}^{k} (1 + x_i + \cdots + x_i^t)^n = \sum_{r_1, \cdots, r_k = 0}^{\infty} f_t(n; r_1, \cdots, r_k) x_1^{r_1} \cdots x_k^{r_k}. \tag{1.12}$$

Then

$$\{n; r_1, \cdots, r_k\}_t = f_t(n; r_1, \cdots, r_k) \qquad (r_j \geqq 0, \quad r_1 + \cdots + r_k \leqq n+1), \tag{1.13}$$

so that

$$\begin{aligned} &\phi_t(x_1, \cdots, x_k) \prod_{i=1}^{k} (1 + x_i + \cdots + x_i^t)^n \\ &\quad = \sum_{r_1 + \cdots + r_k \leqq n} \{n; r_1, \cdots, r_k\}_t x_1^{r_1} \cdots x_k^{r_k} + O_{n+2}(x_1, \cdots, x_n), \end{aligned} \tag{1.14}$$

where

$$O_{n+2}(x_1, \cdots, x_k) = \sum_{r_1 + \cdots + r_k \geqq n+2} a(n; r_1, \cdots, r_k) x_1^{r_1} \cdots x_k^{r_k}.$$

Moreover

$$\{n; r_1, \cdots, r_k\}_t = \sum_{i=1}^{k} \{n, r_i\}_t \prod_{\substack{j=1 \\ j \neq i}}^{k} c_t(n+1, r_j) - (k-1) \prod_{i=1}^{k} c_t(n+1, r_i). \tag{1.15}$$

2. In the present paper we consider the following extension of $\{n; r_1, \cdots, r_k\}$. For brevity put

$$\mathbf{r} = (r_1, \cdots, r_k), \qquad \mathbf{t} = (t_1, \cdots, t_k), \tag{2.1}$$

where the $t_j > 0$. Define the function $\{n; \mathbf{r}; \mathbf{t}\}$ by means of

$$\{n; \mathbf{r}; \mathbf{t}\} = 0 \quad \text{if } r_1 + \cdots + r_k > n \tag{2.2}$$

or if any $r_j < 0$,

$$\{0; \mathbf{0}; \mathbf{t}\} = 1, \tag{2.3}$$

and

$$\{n+1; \mathbf{r}; \mathbf{t}\} = \sum \{n; \mathbf{r} - \boldsymbol{\delta}; \mathbf{t}\}, \tag{2.4}$$

where $\delta = (\delta_1, \cdots, \delta_k)$ and $\delta_j = 0, 1, \cdots, t_j$; thus the right hand side of (2.4) contains $(t_1+1) \cdots (t_k+1)$ terms.

Put

$$\prod_{i=1}^{k} (1 + x_i + \cdots + x_i^{t_i})^n = \sum_{r_1, \cdots, r_k = 0}^{\infty} c(n; \mathbf{r}; \mathbf{t}) x_1^{r_1} \cdots x_k^{r_k}, \tag{2.5}$$

so that

$$c(n;\mathbf{r};\mathbf{t})=\prod_{i=1}^{k} c_{t_i}(n,r_i). \tag{2.6}$$

Also define

$$\phi(\mathbf{x};\mathbf{t})=\left(1-x_1\frac{\partial}{\partial x_1}-\cdots-x_k\frac{\partial}{\partial x_k}\right)\prod_{i=1}^{k}(1+x_i+\cdots+x_i^{t_i}) \tag{2.7}$$

and put

$$\phi(\mathbf{x};\mathbf{t})\prod_{i=1}^{k}(1+x_i+\cdots+x_i^{t_i})^n=\sum_{r_1,\cdots,r_k=0}^{\infty} f(n;\mathbf{r};\mathbf{t})x_1^{r_1}\cdots x_k^{r_k}. \tag{2.8}$$

We show first that

$$f(n;\mathbf{r};\mathbf{t})=0 \quad \text{when } r_1+\cdots+r_k=n+1. \tag{2.9}$$

To prove this result we observe that, by (2.7) and (2.8),

$$\begin{aligned}&\left\{1-\frac{1}{n+1}\left(x_1\frac{\partial}{\partial x_1}+\cdots+x_k\frac{\partial}{\partial x_k}\right)\right\}\prod_{i=1}^{k}(1+x_i+\cdots+x_i^{t_i})^{n+1}\\&=\sum_{r_1,\cdots,r_k=0}^{\infty} f(n;\mathbf{r};\mathbf{t})x_1^{r_1}\cdots x_k^{r_k}.\end{aligned} \tag{2.10}$$

By (2.5), the left member of (2.10) is equal to

$$\left\{1-\frac{1}{n+1}\left(x_1\frac{\partial}{\partial x_1}+\cdots+x_k\frac{\partial}{\partial x_k}\right)\right\}\sum_{r_1,\cdots,r_k=0}^{\infty} c(n+1;\mathbf{r};t)x_1^{r_1}\cdots x_k^{r_k}.$$

Put

$$\sum_{r_1,\cdots,r_k=0}^{\infty} c(n+1;\mathbf{r};\mathbf{t})x_1^{r_1}\cdots x_k^{r_k}=\sum_{r=0}^{\infty} H_r(x_1,\cdots,x_k),$$

where H_r is homogeneous of degree r in $x_1,\cdots,x_k$. Thus, by Euler's theorem on homogeneous functions,

$$\left(x_1\frac{\partial}{\partial x_1}+\cdots+x_k\frac{\partial}{\partial x_k}\right)H_{n+1}=(n+1)H_{n+1}$$

and (2.9) follows at once.

Next, by (2.8),

$$\begin{aligned}&\sum_{r_1,\cdots,r_k=0}^{\infty} f(n+1;\mathbf{r};\mathbf{t})x_1^{r_1}\cdots x_k^{r_k}\\&\qquad=\prod_{i=1}^{k}(1+x_i+\cdots+x_i^{t_i})\cdot\sum_{r_1,\cdots,r_k=0}^{\infty} f(n;\mathbf{r};\mathbf{t})x_1^{r_1}\cdots x_k^{r_k},\end{aligned}$$

which gives

$$f(n+1;\mathbf{r};\mathbf{t})=\sum_{\delta_j=0}^{t_j} f(n;\mathbf{r}-\boldsymbol{\delta};\mathbf{t}) \tag{2.11}$$

for all nonnegative $r_1, \cdots, r_k$. It is understood that

$$f(\mathbf{n}; \mathbf{r}; \mathbf{t}) = 0 \tag{2.12}$$

if any $r_i < 0$.

Comparing (2.9), (2.11) and (2.12) with (2.2), (2.3) and (2.4) we infer that

$$\{n; \mathbf{r}; t\} = f(n; \mathbf{r}; \mathbf{t}) \qquad (r_1 + \cdots + r_k \leqq n+1). \tag{2.13}$$

We may therefore state the following theorem.

THEOREM 1. *The function* $\{n; \mathbf{r}; \mathbf{t}\}$ *defined by* (2.2), (2.3) *and* (2.4) *satisfies*

$$\begin{aligned} &\phi(\mathbf{x}; \mathbf{t}) \prod_{i=1}^{k} (1 + x_i + \cdots + x_i^{t_i})^n \\ &\qquad = \sum_{r_1 + \cdots + r_k \leqq n} \{n; \mathbf{r}; \mathbf{t}\} x_1^{r_1} \cdots x_k^{r_k} + O_{n+2}(x_1, \cdots, x_k), \end{aligned} \tag{2.14}$$

where $\phi(\mathbf{x}; \mathbf{t})$ *is defined by* (2.7) *and*

$$O_{n+2}(x_1, \cdots, x_k) = \sum_{r_1 + \cdots + r_k \geqq n+2} a(n; \mathbf{r}) x_1^{r_1} \cdots x_k^{r_k},$$

where the coefficients $a(n; r)$ *are independent of* $x_1, \cdots, x_k$.

3. It follows from (2.7) and (1.6) that

$$\phi(\mathbf{x}; \mathbf{t}) = \sum_{i=1}^{k} \phi_{t_i}(x_i) \prod_{\substack{j=1 \\ j \neq i}}^{k} (1 + x_j + \cdots + x_j^{t_j}) - (k-1) \prod_{i=1}^{k} (1 + x_i + \cdots + x_i^{t_i}). \tag{3.1}$$

Thus

$$\begin{aligned} &\phi(\mathbf{x}; \mathbf{t}) \prod_{i=1}^{k} (1 + x_i + \cdots + x_i^{t_i})^n \\ &\qquad = \sum_{i=1}^{k} \phi_{t_i}(x_i)(1 + x_i + \cdots + x_i^{t_i})^n \prod_{\substack{j=1 \\ j \neq i}}^{k} (1 + x_j + \cdots + x_j^{t_j})^{n+1} \\ &\qquad\quad - (k-1) \prod_{i=1}^{k} (1 + x_i + \cdots + x_i^{t_i})^{n+1}. \end{aligned} \tag{3.2}$$

Hence the left hand side of (3.2) is equal to

$$\begin{aligned} &\sum_{i=1}^{k} \left\{ \sum_{r_i=0}^{n} \{n, r_i\}_{t_i} x_i^{t_i} + O(x_i^{n+2}) \right\} \prod_{\substack{j=1 \\ j \neq i}}^{\infty} c_{t_j}(n+1, r_j) x_j^{r_j} \\ &\qquad - (k+1) \prod_{i=1}^{k} \sum_{r_i=0}^{\infty} c_{t_i}(n+1, r_i) x_i^{r_i}. \end{aligned} \tag{3.3}$$

Making use of (2.14), (3.2) and (3.3) we can express $\{n; \mathbf{r}; \mathbf{t}\}$ in terms of $\{n; r\}_t$ and $c_t(n+1, r)$.

THEOREM 2. *The function* $\{n; \mathbf{r}; \mathbf{t}\}$ *is expressed in terms of* $\{n; r\}_t$ *and* $c_t(n+1, r\}$ *by means of*

$$\{n; \mathbf{r}; \mathbf{t}\} = \sum_{i=1}^{k} \{n, r_i\}_{t_i} \prod_{\substack{j=1 \\ j \neq i}}^{k} c_{t_j}(n+1, r_j) - (k-1) \prod_{i=1}^{k} c_{t_i}(n+1, r_i). \tag{3.4}$$

4. As an application of the function $\{n; \mathbf{r}, \mathbf{t}\}$ we consider $k \times n$ arrays of nonnegative integers

$$A = (a_{ij}) \qquad (1 \leqq i \leqq k; \quad 1 \leqq j \leqq n) \tag{4.1}$$

such that

$$0 \leqq a_{i,j+1} - a_{i,j} \leqq t_i \qquad (1 \leqq i \leqq k; \quad 1 \leqq j < n) \tag{4.2}$$

and

$$\sum_{i=1}^{k} a_{ij} \leqq j \qquad (1 \leqq j \leqq n). \tag{4.3}$$

For example, for $k = 1$, we have sequences $(a_1, a_2, \cdots, a_n)$ of nonnegative integers such that

$$0 \leqq a_{j+1} - a_j \leqq t, \qquad a_j \leqq j \qquad (1 \leqq j \leqq n). \tag{4.4}$$

As noted in [1], for $t = \infty$, the number of sequences satisfying (4.4) is a ballot number [2, III, Ch. 5].

Now let $N(n; \mathbf{r}, \mathbf{t})$ denote the number of arrays (4.1) that satisfy (4.2) and (4.3) and in addition

$$a_{in} = r_i \qquad (1 \leqq i \leqq k). \tag{4.5}$$

It follows from the definition that $N(n; \mathbf{r}, \mathbf{t})$ satisfies the following relations:

$$N(n; \mathbf{r}, \mathbf{t}) = 0 \quad \text{if } \sum_{i=1}^{k} r_i > n \tag{4.6}$$

or if any $r_i < 0$,

$$N(n+1; \mathbf{r}, \mathbf{t}) = \sum N(n; \mathbf{r} - \boldsymbol{\delta}; \mathbf{t}), \tag{4.7}$$

where $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_k)$ and the summation is over $\delta_i = 0, 1, \cdots, t_i$ $(1 \leqq i \leqq k)$. In addition we define

$$N(0; 0; \mathbf{t}) = 1 \tag{4.8}$$

so that (4.7) holds for all $n \geqq 0$.

Comparing (4.6), (4.7) and (4.8) with (2.2), (2.3), (2.4), we conclude that

$$N(n; \mathbf{r}, \mathbf{t}) = \{n; \mathbf{r}, \mathbf{t}\}. \tag{4.9}$$

Thus $\{n; \mathbf{r}, \mathbf{t}\}$ furnishes a generalization of ballot numbers in several directions.

A generalization of $N(n; \mathbf{r}, \mathbf{t})$ can be obtained by replacing (4.3) by

$$\sum_{i=1}^{k} a_{ij} \leqq j + r \qquad (1 \leqq j \leqq n), \tag{4.10}$$

where r is some fixed nonnegative integers. A generalization of a different kind suggested by [2] is the following. Let $(N; \mathbf{r}, \mathbf{t}, q)$ denote the sum

$$\sum q^{|A|}, \qquad |A| = \sum_{i=1}^{k} \sum_{j=1}^{n} a_{ij}, \tag{4.11}$$

where the first summation is over all arrays (4.1) that satisfy (4.2), (4.3) and (4.5), and q is independent of n.

We shall not investigate these generalizations in the present paper.

REFERENCES

[1] L. CARLITZ, *Solution of certain recurrences*, SIAM J. Appl. Math., 17 (1969), pp. 251–259.

[2] L. CARLITZ AND J. RIORDAN, *Two element lattice permutation numbers and their q-generalization*, Duke Math. J., 31 (1964), pp. 371–388.

[3] P. A. MACMAHON, *Combinatorial Analysis*, vol. 1, Cambridge University Press, Cambridge, England, 1915.

[4] T. V. NARAYANA AND V. K. ROHATGI, *Further analogues of the multinomial theorem*, J. Indian Statist. Assoc., 5 (1967), pp. 83–89.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0002 $01.00/0

ON A CLASS OF NONLINEAR INTEGRAL EQUATIONS ARISING IN TRANSPORT THEORY*

G. A. HIVELY†

Abstract. For a class of kernels $k(x, t)$, conditions are given for the existence and uniqueness of a solution in the unit ball of $L^1(\mu)$ for the nonlinear integral equation

$$u(x) = \psi(x) + u(x) \int k(x, t)u(t)\, d\mu(t).$$

Equations of this type arise in the theories of radiative transfer, neutron transport and in the kinetic theory of gases.

1. Introduction. In the theories of radiative transfer (cf. [2], [4]) and neutron transport (cf. [3], [7]) an important role is played by nonlinear integral equations of the form

$$H(x) = 1 + xH(x) \int_0^1 \frac{\psi(t)H(t)}{x+t}\, dt, \tag{1.1}$$

where ψ is a given function on [0, 1]. Multiplying this equation by ψ one obtains the equation

$$u(x) = \psi(x) + u(x) \int_0^1 \frac{x}{x+t} u(t)\, dt, \tag{1.2}$$

for the function $u(x) = \psi(x)H(x)$. Equations similar to (1.1) and (1.2) for functions on the half-line $[0, \infty)$ arise in the kinetic theory of gases (cf. [6]).

In [1], Bittoni, Casadei and Lorenzutta have studied equation (1.2) in the Banach space $L^1(0, 1)$ from the point of view of contraction mappings. Their main result is that if $\|\psi\|_1 < 1/2$ and $\psi \geqq 0$ then (1.2) admits a unique solution u in a certain sphere of $L^1(0, 1)$. In attempting to extend and improve this result we have found (a) that when $\|\psi\|_1 < 1/2$ the assumption $\psi \geqq 0$ is superfluous, (b) that when $\psi \geqq 0$ the assumption $\|\psi\|_1 \leqq 1/2$ is sufficient (so that the important conservative case is included) and (c) that the arguments which establish these results are valid for a more general class of equations.

Let (X, μ) be a σ-finite positive measure space with $\mu \neq 0$ and let $k(x, t)$ be a measurable function on the product measure space $(X \times X, \mu \times \mu)$ satisfying

(i) $0 < k(x, t) < 1$, $x, t \in X$,

(ii) $k(x, t) + k(t, x) = 1$, $x, t \in X$.

Let $L^1 = L^1(X, \mu)$, $L^\infty = L^\infty(X, \mu)$ and let K be the integral operator on L^1 with kernel $k(x, t)$,

$$Ku(x) = \int k(x, t)u(t)\, d\mu(t), \qquad u \in L^1.$$

Clearly K is a bounded linear operator from L^1 into L^∞ with $\|K\|_{1,\infty} \leqq 1$. In this paper we shall consider the problem of solving the nonlinear equation

$$u = \psi + uKu, \qquad u \in L^1, \tag{1.3}$$

* Received by the editors August 3, 1976.

† Department of Computer Science, University of Kentucky, Lexington, Kentucky 40506. This research was carried out during the author's participation in the Summer Graduate Research Program in the Applied Mathematics Division of Argonne National Laboratory and was supported by U.S. ERDA.

where $\psi \in L^1$ is given. We note that equation (1.2) is of this form, where the kernel $k(x, t)$ is $x(x+t)^{-1}$ and satisfies conditions (i) and (ii) on any subinterval X of $(0, \infty)$. Our principal results are that equation (1.3) has a unique solution in the closed unit ball of L^1 for each $\psi \in L^1$ satisfying either of the conditions (a) $\|\psi\|_1 < 1/2$ or (b) $\psi \geqq 0$ and $\|\psi\|_1 \leqq 1/2$, and that under either condition the solution depends continuously upon ψ and may be obtained by an iterative procedure.

We remark that it is not until § 4 that we require the full strength of assumption (i). For the results of §§ 2 and 3, it is enough that $0 \leqq k(x, t) \leqq 1$.

2. The case $\|\psi\|_1 < 1/2$. The starting point for the analysis of [1] is the observation that equation (1.3) has the form

$$u = \psi + A(u, u), \qquad u \in L^1, \tag{2.1}$$

where A is the bilinear mapping on L^1 defined by $A(u, v) = uKv$ for $u, v \in L^1$ and that the solutions of (2.1) are precisely the fixed points for the mapping T_ψ on L^1 defined by $T_\psi u = \psi + A(u, u)$ for $u \in L^1$.

In general, recall that a bilinear mapping B on a Banach space Y is said to be bounded if its norm, $\|B\| = \sup\{\|B(x, y)\|: \|x\|, \|y\| \leqq 1\}$, is finite. If B is a bilinear mapping on Y, let B^* be the bilinear mapping defined by $B^*(x, y) = B(y, x)$. If B is bounded then so are B^* and $B + B^*$. The significance of $B + B^*$ lies in the fact that if B is bounded then the bounded linear mapping $(B + B^*)(x, \cdot)$ is the Fréchet derivative at $x \in Y$ of the mapping $y \mapsto B(y, y)$. Although the mean value theorem does not hold in general for vector valued functions, it turns out that in the present context it does hold. Indeed, for $u, v \in Y$ one has

$$B(u, u) - B(v, v) = (B + B^*)\left(\frac{u+v}{2}, u - v\right). \tag{2.2}$$

Using this simple formula we are able to obtain the following result without an implicit use of the Hahn-Banach theorem (see the reference to [4] in [1]).

THEOREM 1. *Let Y be a Banach space and let B be a bounded bilinear mapping on Y such that $\|B + B^*\| > 0$. Let*

$$S_1 = \{u \in Y: \|u\| \leqq \|B + B^*\|^{-1}\},$$
$$S_2^0 = \{y \in Y: \|y\| < \tfrac{1}{2}\|B + B^*\|^{-1}\}$$

and for each $y \in Y$ let the mapping T_y on Y be defined by

$$T_y u = y + B(u, u), \qquad u \in Y.$$

Then for each $y \in S_2^0$ the mapping T_y maps S_1 into itself, has a unique fixed point $u_y \in S_1$ and for each $u \in S_1$ the sequence $\{T_y^n u\}$ converges to u_y. Moreover, the mapping $y \mapsto u_y$ of S_2^0 into S_1 is continuous.

Proof. Let $y \in S_2^0$ and let

$$S_y = \{u \in Y: \|u - y\| \leqq \tfrac{1}{2}\|B + B^*\|^{-1}\}.$$

Then $S_y \subseteq \lambda S_1$, where $\lambda = \frac{1}{2} + \|y\| \|B + B^*\| < 1$. If $u \in S_1$ then we have $\|T_y u - y\| = \|B(u, u)\| = \frac{1}{2}\|(B + B^*)(u, u)\| \leqq \frac{1}{2}\|B + B^*\|^{-1}$, so that $T_y(S_1) \subseteq S_y$. Thus we have both $T_y(S_1) \subseteq S_1$ and $T_y(S_y) \subseteq S_y$. Using (2.2) we find that if $u, v \in S_y \subseteq \lambda S_1$ then $\|T_y u - T_y v\| \leqq \lambda \|u - v\|$. Thus T_y maps S_y into itself and is a strict Lipschitz contraction on S_y and so the existence and uniqueness of a fixed point $u_y \in S_y$ follows. But $T_y(S_1) \subseteq S_y \subseteq$

S_1 and so u_y exists and is unique in S_1. Also, if $u \in S_1$, then $T_y u \in S_y$ and so the sequence $\{T_y^n u\}$ will converge to u_y.

In order to prove the last assertion it suffices to show that the mapping $y \mapsto u_y$ is continuous on each εS_2^0 for $0<\varepsilon<1$. So let $0<\varepsilon<1$ be given. From the preceding argument we see that if $y \in \varepsilon S_2^0$ then $u_y \in S_y \subseteq \lambda S_1$, where $\lambda = \frac{1}{2}+\frac{1}{2}\varepsilon < 1$. If $x, y \in \varepsilon S_2^0$ then from (2.2) and the convexity of λS_1 we have

$$\|u_x - u_y\| \leqq \|x-y\| + \lambda \|u_x - u_y\|$$

so that

$$\|u_x - u_y\| \leqq (1-\lambda)^{-1} \|x-y\|.$$

Thus the mapping $y \mapsto u_y$ is even Lipschitz continuous on εS_2^0. Q.E.D.

In order to apply Theorem 1 to equation (2.1) we must determine $\|A+A^*\|$. From the hypotheses (i) and (ii) on $k(x, t)$ we have that for $u, v \in L^1$.

$$\begin{aligned} &\|(A+A^*)(u, v)\|_1 \\ &\qquad \leqq \int\int k(x,t)\{|u(x)|\,|v(t)|+|u(t)|\,|v(x)|\}\,d\mu(t)\,d\mu(x) \\ &\qquad = \int\int \{k(x,t)+k(t,x)\}|u(x)|\,|v(t)|\,du(t)\,d\mu(x) \\ &\qquad = \|u\|_1\|v\|_1, \end{aligned}$$

showing that $\|A+A^*\| \leqq 1$. Taking $u, v \geqq 0$ we find that $\|A+A^*\|=1$. This fact together with Theorem 1 yields

THEOREM 2. *Let* $S_2^0 = \{u \in L^1 : \|u\|_1 < 1/2\}$. *Then for each* $\psi \in S_2^0$ (1.3) *admits a unique solution* u_ψ *in the closed unit ball* S_1 *of* L^1, T_ψ *maps* S_1 *into itself and for each* $u \in S_1$ *the sequence* $\{T_\psi^n u\}$ *converges to* u_ψ *in* L^1. *Moreover, the mapping* $\psi \mapsto u_\psi$ *is continuous on* S_2^0.

3. The case $\psi \geqq 0$. We now turn to the problem of solving equation (1.3) in the case where $\psi \geqq 0$ a.e. and $\|\psi\|_1 \leqq 1/2$. We require two preliminary results, the first of which is quite standard (cf. [2, Thm. 12.1]).

LEMMA 1. *Let* $u, \psi \in L^1$ *and let u satisfy equation* (1.3). *Then* $U^2 - 2U + 2\Psi = 0$, *where U and* Ψ *are the integrals of u and* ψ *respectively.*

Proof. Integrating the equation $u = \psi + uKu$ and using the identity

$$k(x,t) = \tfrac{1}{2}\{1 + k(x,t) - k(t,x)\}$$

we have

$$U = \Psi + \tfrac{1}{2}U^2 + \tfrac{1}{2}\int\int \{k(x,t)-k(t,x)\}u(x)u(t)\,d\mu(t)\,d\mu(x)$$

and the last integral is clearly zero. Q.E.D.

LEMMA 2. *If* $\phi, \psi \in L^1$ *with* $0 \leqq \phi \leqq \psi$ *and if* $u, v \in L^1$ *with* $0 \leqq u \leqq v$ *then*

$$0 \leqq T_\phi^n u \leqq T_\psi^n v, \qquad n = 0, 1, \cdots.$$

Proof. It suffices to show that $0 \leqq T_\phi u \leqq T_\psi v$. For this, note that if $u_1, u_2 \geqq 0$ then $u_1 K u_2 \geqq 0$. Thus

$$T_\phi u = \phi + uKu \geqq \phi \geqq 0$$

and

$$\begin{aligned} T_\psi v - T_\phi u &= (\psi - \phi) + vKv - uKu \\ &= (\psi - \phi) + (v - u)Kv + uK(v - u) \\ &\geqq 0. \end{aligned} \qquad \text{Q.E.D.}$$

THEOREM 3. *Let* $S_2^+ = \{u \in L^1 : \|u\|_1 \leqq \frac{1}{2}, u \geqq 0\}$. *Then for each* $\psi \in S_2^+$ *equation* (1.3) *admits a unique solution* u_ψ *in the closed unit ball* S_1 *of* L^1, $u_\psi \geqq 0$, T_ψ *maps* S_1 *into itself and for any* $u \in S_1$ *with* $u \geqq 0$ *the sequence* $\{T_\psi^n u\}$ *converges to* u_ψ *in* L^1. *Moreover, the mapping* $\psi \mapsto u_\psi$ *is uniformly continuous on* S_2^+.

Proof. Let $\psi \in S_2^+$. Since $\|\psi\|_1 \leqq 1/2$ it follows, as in the proof of Theorem 1, that T_ψ maps S_1 into itself. Setting $u_n = T_\psi^n 0$ we have $0 \leqq u_0 \leqq u_1$ and so, by Lemma 2, it follows that $\{u_n\}$ is an increasing sequence in L^1. Since this sequence is contained in S_1 it follows by the monotone convergence theorem that it converges in L^1 to some $u_\psi \in S_1$ with $u_\psi \geqq 0$. Since $u_{n+1} = \psi + u_n K u_n$ and $u_n \to u_\psi$ in L^1 it follows that $Ku_n \to Ku_\psi$ in L^∞ and $u_\psi = \psi + u_\psi K u_\psi$. This proves the existence of a solution u_ψ of (1.3) with $u_\psi \in S_1$ and $u_\psi \geqq 0$.

If $\|\psi\|_1 < 1/2$ then the uniqueness of u_ψ in S_1 follows from Theorem 2. So suppose that $\|\psi\|_1 = 1/2$ and that u is a solution of (1.3) with $u \in S_1$. We shall show that $u = u_\psi$ where u_ψ is the particular solution obtained in the preceding paragraph. Since u and u_ψ are solutions of (1.3), Lemma 1 implies that $\int u(x)\,d\mu(x) = \int u_\psi(x)\,d\mu(x) = 1$, so that, since $\|u\|_1 \leqq 1$, we must have $u \geqq 0$. From Lemma 2 we have that $u_n = T_\psi^n 0 \leqq T_\psi^n u = u$ for each $n \geqq 0$. Since $u_n \to u_\psi$ in L^1 it follows that $u_\psi \leqq u$. But then $\|u - u_\psi\|_1 = \int u(x)\,d\mu(x) - \int u_\psi(x)\,d\mu(x) = 0$ so that $u = u_\psi$. This shows the uniqueness of u_ψ in S_1.

Now let $v \in S_1$ with $v \geqq 0$. We must show that the sequence $\{v_n\} = \{T_\psi^n v\}$ converges to u_ψ in L^1. If $\|\psi\|_1 < \frac{1}{2}$ then this follows from Theorem 2 and so we may assume that $\|\psi\|_1 = \frac{1}{2}$. In this case we have, as above, that $u_n = T_\psi^n 0 \leqq T_\psi^n v = v_n$, $u_n \to u_\psi$ in L^1 and $\int u_\psi(x)\,d\mu(x) = 1$. Since $v_n \in S_1$ it follows that $(v_n - u_n) \to 0$ in L^1 and therefore $v_n = (v_n - u_n) + u_n \to u_\psi$ in L^1.

If $\phi, \psi \in S_2^+$ with $\phi \leqq \psi$ then from the result just established we have $T_\phi^n \phi \to u_\phi$ in L^1 and $T_\psi^n \psi \to u_\psi$ in L^1 and, from Lemma 2, $0 \leqq T_\phi^n \phi \leqq T_\psi^n \psi$ so that we must have $0 \leqq u_\phi \leqq u_\psi$. Let Φ, Ψ, U_ϕ and U_ψ denote the integrals of ϕ, ψ, u_ϕ and u_ψ respectively. Then using Lemma 1 we have

$$\begin{aligned} \|u_\psi - u_\phi\|_1 = U_\psi - U_\phi &= (1 - 2\Phi)^{1/2} - (1 - 2\Psi)^{1/2} \\ &\leqq \sqrt{2}(\Psi - \Phi)^{1/2} = \sqrt{2}\|\psi - \phi\|_1^{1/2}. \end{aligned}$$

Now let $\psi_1, \psi_2 \in S_2^+$ and define $\phi = \psi_1 \wedge \psi_2$ so that $\phi \in S_2^+$, $0 \leqq \phi \leqq \psi_1, \psi_2$ and $\|\psi_1 - \phi\|_1 + \|\psi_2 - \phi\|_1 = \|\psi_1 - \psi_2\|_1$. If u_1, u_2 and u are the solutions of (1.3) in S_1 for ψ_1, ψ_2 and ϕ respectively, then using the inequality established in the previous paragraph we have

$$\begin{aligned} \|u_1 - u_2\|_1 &\leqq \|u_1 - u\|_1 + \|u_2 - u\|_1 \\ &\leqq \sqrt{2}\|\psi_1 - \phi\|_1^{1/2} + \sqrt{2}\|\psi_2 - \phi\|_1^{1/2} \\ &\leqq 2\sqrt{2}\|\psi_1 - \psi_2\|_1^{1/2}, \end{aligned}$$

which shows that the mapping $\psi \mapsto u_\psi$ is uniformly continuous on S_2^+. Q.E.D.

4. Continuous kernels. In this final section we consider equation (1.3) in the presence of a topology on X (cf. [1, Theorem 2]). For the sake of simplicity we shall assume that X is equipped with a metric topology and that μ is a Borel measure on X

with $\operatorname{supp}(\mu)=X$ (i.e., each nonempty open set has positive measure). In addition to the previous assumptions (i) and (ii), we assume that the kernel satisfies

(iii) $k(x, t)$ is continuous on $X\times X$.

In the next result we use, for the first time, the full strength of assumption (i).

LEMMA 3. *For each $u\in L^1$, Ku is continuous on X and, if $u\neq 0$, then $|Ku(x)|<\|u\|_1$ for each $x\in X$.*

Proof. If $x\in X$ and $\{x_n\}$ is a sequence in X converging to x, then

$$Ku(x_n)-Ku(x)=\int \{k(x_n, t)-k(x, t)\}u(t)\,d\mu(t),$$

and the integral converges to zero by (i), (iii) and the dominated convergence theorem. Thus Ku is continuous on X. If $u\neq 0$ and $x\in X$, then

$$|Ku(x)|\leqq \int k(x, t)|u(t)|d\mu(t)<\|u\|_1,$$

where the last inequality follows from assumption (i). Q.E.D.

THEOREM 4. *If $u\in S_1$ and u is a solution of equation* (1.3), *then u is continuous. wherever ψ is continuous. If ψ is continuous, $\psi\geqq 0$ and $\|\psi\|_1\leqq\frac{1}{2}$, then the unique solution u_ψ in S_1 of* (1.3) *is continuous and if u is continuous, with $0\leqq u\leqq\psi$, then the sequence $\{T_\psi^n u\}$ consists of continuous functions converging monotonically and almost uniformly to u_ψ.*

Proof. Let $u\in S_1$ be a solution of equation (1.3). Then $\psi=u(1-Ku)$ and, by Lemma 3, $1-Ku$ is continuous and nonvanishing on X. Thus u must be continuous wherever ψ is continuous.

Now let ψ be continuous, $\psi\geqq 0$ and $\|\psi\|_1\leqq\frac{1}{2}$. By Theorem 3, equation (1.3) has a unique solution $u_\psi\in S_1$, which, since ψ is continuous, must also be continuous. If u is continuous and $0\leqq u\leqq\psi$, then each $T_\psi^n u$ is clearly continuous and, by Theorem 3, $T_\psi^n u\to u_\psi$ in L^1. Since $T_\psi u=\psi+uKu\geqq\psi\geqq u$, it follows from Lemma 2 that $\{T_\psi^N u\}$ is an increasing sequence. Since supp $(\mu)=X$, we must have $T_\psi^n u\leqq u_\psi$ for each $n\geqq 0$. In order to show that the sequence $\{T_\psi^n u\}$ converges almost uniformly to u_ψ, it suffices, by Dini's theorem, to show that the sequence converges pointwise to u_ψ on X. But letting $u_n=T_\psi^n u$ we have

$$\begin{aligned}(1-Ku_\psi)(u_\psi-u_{n+1})&=u_{n+1}Ku_\psi+(1-Ku_\psi)u_\psi-u_{n+1}\\&=u_{n+1}Ku_\psi+\psi-u_{n+1}\\&=u_{n+1}Ku_\psi-u_nKu_n\\&=(u_{n+1}-u_n)Ku_\psi+u_nK(u_\psi-u_n).\end{aligned}$$

For each $x\in X$ we have $1-Ku_\psi(x)\neq 0$ by Lemma 2, $K(u_\psi-u_n)\to 0$ uniformly since $u_n\to u_\psi$ in L^1, $\{u_n(x)\}$ is bounded and $u_{n+1}(x)-u_n(x)\to 0$ since $\{u_n(x)\}$ is monotone increasing and bounded above by $u_\psi(x)$. Thus the sequence $\{u_n\}=\{T_\psi^n u\}$ converges pointwise, and hence almost uniformly, to u_ψ. Q.E.D.

Acknowledgment. The author wishes to thank Dr. Hans G. Kaper for his generous help and encouragement.

REFERENCES

[1] E. BITTONI, G. CASADEI AND S. LORENZUTTA, *Una applicazione del principio di contrazione alla equazione di Ambartsumian–Chandrasekhar*, Boll. Un. Mat. Ital., 2(1969), pp. 535–541.

[2] I. W. BUSBRIDGE, *The Mathematics of Radiative Transfer*, Cambridge Tracts in Mathematical Physics No. 50, Cambridge University Press, Cambridge, England, 1960.

[3] K. M. CASE, P. F. ZWEIFEL, *Linear Transport Theory*, Addison-Wesley, Reading, MA, 1967.

[4] S. CHANDRASEKHAR, *Radiative Transfer*, Oxford University Press, London, 1950.

[5] L. V. KANTOROVICH AND G. P. AKILOV, *Functional Analysis in Normed Spaces*, Pergamon Press, London, 1964.

[6] H. G. KAPER, *A Constructive Approach to the Solution of a Class of Boundary Value Problems of Mixed Type in the Kinetic Theory of Gases*, J. Math. Anal. Appl, to appear.

[7] N. J. MCCORMICK AND I. KUŠČER, *Singular eigenfunction expansions in neutron transport theory*, Advances in Nuclear Science and Technology, Vol. 7, E. J. Henley and J. Lewins, eds., Academic Press, New York, 1973, pp. 181–282.

SIAM J. MATH ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0003 $01.00/0

AN EXISTENCE RESULT FOR A NONLINEAR VOLTERRA INTEGRAL EQUATION IN A HILBERT SPACE*

GUSTAF GRIPENBERG†

Abstract. We study equations of the form

$$u(t)+\int_0^t a(t-s)gu(s)\,ds \ni f(t), \qquad t\geqq 0$$

on a real Hilbert space H. The unknown function is u and a, g, f are given. It is assumed that the kernel a is operator-valued (real-valued as a special case) and g is an arbitrary maximal monotone operator in H. The method can also be applied to time-varying nonlinearity. We prove an existence and uniqueness result that extends earlier results by Londen and Barbu. Finally an application is given.

1. Introduction. We consider existence and uniqueness of solutions of the nonlinear Volterra equation

$$u(t)+\int_0^t a(t-s)gu(s)\,ds \ni f(t), \qquad t\geqq 0, \tag{1.1}$$

where a, g and f are given and u is the unknown taking values in a real Hilbert space H. The kernel a is a real-valued function and the nonhomogeneous term f takes $[0,\infty)$ into H. The mapping g is a nonlinear monotone (in general multi-valued) operator, having its domain Dg and range Rg in H. The integral in (1.1) is to be considered as a Bochner integral.

We say that a function $u\colon [0,T]\to H$ is a solution of (1.1) on $[0,T]$ if the following conditions hold:

$$u\in L^2(0,T;H), \tag{1.2}$$

$$u(t)\in Dg \quad \text{a.e.} \quad \text{on } [0,T], \tag{1.3}$$

$$\exists w\in L^2(0,T;H), \tag{1.4}$$

such that

$$w(t)\in gu(t) \quad \text{a.e.} \quad \text{on } [0,T], \tag{1.5}$$

and

$$u(t)+\int_0^t a(t-s)w(s)\,ds = f(t) \quad \text{a.e.} \quad \text{on } [0,T]. \tag{1.6}$$

A function $u\colon [0,\infty)\to H$ is said to be a solution on $[0,\infty)$ if it is a solution on $[0,T]$ for any $T>0$. We will also consider the more general equation

$$u(t)+\int_0^t A(t-s)g(s)u(s)\,ds \ni f(t), \qquad t\geqq 0, \tag{1.7}$$

with operator-valued kernel and time-varying nonlinearity. Thus we assume that the kernel A takes $[0,\infty)$ into $L(H)$ (the Banach space of bounded linear operators on H) and that $g(s)$ is a nonlinear (multi-valued) mapping with domain $Dg(s)$ and range $Rg(s)$ in H for a.e. $s\geqq 0$.

* Received by the editors October 14, 1976, and in revised form January 20, 1977.
† Institute of Mathematics, Helsinki University of Technology, Otaniemi, Finland.

We say that a function $u: [0, T] \to H$ is a solution of (1.7) on $[0, T]$ if in addition to (1.2) and (1.4) the following conditions hold:

$$u(t) \in Dg(t) \quad \text{a.e.} \quad \text{on } [0, T], \tag{1.8}$$

$$w(t) \in g(t)u(t) \quad \text{a.e.} \quad \text{on } [0, T], \tag{1.9}$$

and

$$u(t) + \int_0^t A(t-s)w(s)\,ds = f(t) \quad \text{a.e.} \quad \text{on } [0, T]. \tag{1.10}$$

Again a function $u: [0, \infty) \to H$ is said to be a solution on $[0, \infty)$ if it is a solution on $[0, T]$ for every $T > 0$.

Our results are formulated in Theorems 1 and 2 of § 2. These results concern existence and uniqueness of solutions of equations (1.1) and (1.7) respectively. In § 2 we also give some comments including comparisons to earlier related results. In § 3 we prove Theorem 2. (Theorem 1 follows as a special case.) In § 4 we give some examples. Our key assumption on the nonlinear functions $g(s)$ is that they should be maximal monotone mappings. We do not however require these mappings to be subdifferentials of convex functions.

Our notational habit follows that of [4]. Thus for example $(\cdot, \cdot)$ denotes the scalar product in H and $|\cdot|$ denotes the norm in H and $L(H)$. We say that a function v is of essentially bounded variation if it is a.e. equal to a function of bounded variation.

2. Statement of results.

THEOREM 1. *Assume*

(2.1) $a \in W^{1,1}(0, T; R), \quad \forall T > 0,$

(2.2) a' *is of essentially bounded variation on* $[0, T], \quad \forall T > 0,$

(2.3) $a(0) > 0,$

(2.4) g *is a maximal monotone operator in* H,

(2.5) $f \in W^{1,1}(0, T; H), \quad \forall T > 0,$

(2.6) f' *is of essentially bounded variation on* $[0, T], \quad \forall T > 0,$

(2.7) $f(0) \in Dg.$

Then there exists a unique solution u *of* (1.1) *on* $[0, \infty)$ *satisfying*

(2.8) u *is Lipschitz continuous on* $[0, T], \quad \forall T > 0,$

(2.9) $w \in L^\infty(0, T; H), \quad \forall T > 0,$

and such that (1.6) *holds for every* $t \geqq 0$.

In Theorem 1 in [7] Londen uses (2.1) and (2.3) but assumes (2.2) to hold only for some $T_0 > 0$ and instead of (2.5), (2.6) the condition $f \in W^{1,2}(0, T; H)$, $\forall T > 0$. The main difference between our result and that of Londen is however that in the latter g is assumed to be the subdifferential of a proper, lower semicontinuous convex function $\varphi: H \to (-\infty; \infty]$. Our assumption (2.4) is obviously much weaker. But note that we then have to assume $f(0) \in Dg$ whereas Theorem 1 in [7] uses only $f(0) \in D\varphi$.

In [1] Barbu has studied (1.1) under the hypothesis that g is a subdifferential. In addition he analyzes the case when g satisfies certain continuity and boundedness conditions and is not necessarily a subdifferential. He uses kernels of positive type, the main advantage being that one can allow $a(0+) = +\infty$. In [3] Barbu has also considered certain operator-valued kernels.

If the kernel $a(t)\equiv 1$, $t\geqq 0$ then (1.1) formally reduces to the evolution equation

$$\frac{du}{dt}+gu(t)\ni f'(t), \qquad u(0)=f(0).$$

This equation has been extensively studied, see for example [4].

It is quite clear that Theorem 1 is a special case of the following:

THEOREM 2. *Assume* (2.5), (2.6) *and*

(2.10) $A\in W^{1,1}(0,T;L(H)), \qquad \forall T>0,$

(2.11) A' *is of essentially bounded variation on* $[0,T]$, $\qquad \forall T>0,$

(2.12) $\exists T_1>0$ *such that* $A'(t)$ *is symmetric for* a.e. $t\in[0,T_1]$,

(2.13) $A(0)=a_0I, \qquad a_0>0,$

(2.14) $\exists$ *a measurable set* $E\subset[0,\infty)$ *such that* $m\{[0,\infty)\backslash E\}=0$, $0\in E$, *and* $t\in E\Rightarrow g(t)$ *is a maximal monotone operator in* H,

(2.15) $\exists$ *a function* $e:[0,\infty)\to R$, $\exists T_2>0$, $\exists\lambda_0>0$ *and* $\exists$ *a constant* L *such that* var $(e;[t,t+T_2])\leqq q<1 \quad \forall t\geqq 0$, *and such that*

(2.16) $|g_\lambda(t)x-g_\lambda(s)x|\leqq|e(t)-e(s)|(L(|x|+1)+|g_\lambda(s)x|) \quad \forall x\in H,$
$\forall\lambda\in(0,\lambda_0], \qquad \forall s,t\in E, \qquad s\leqq t,$

(2.17) $f(0)\in Dg(0)$.

Then there exists a unique solution u *of* (1.7) *satisfying*

(2.18) u *is Lipschitz continuous on* $[0,T]$, $\quad \forall T>0,$

(2.19) $w\in L^\infty(0,T;H)$, $\quad \forall T>0$ *and such that* (1.10) *holds for every* $t\geqq 0$.

If $A(t)\equiv I$, $t\geqq 0$, then (1.7) formally reduces to

$$\frac{du}{dt}+g(t)u(t)\ni f'(t), \qquad u(0)=f(0).$$

This equation has been studied among others by Crandall and Pazy [5] and Evans [6]. The assumptions they use on the t-dependence of $g(t)$ are similar to the hypothesis used here but on the whole also weaker. It is to be observed that they study the equation in a general Banach space.

3. Proof of Theorem 2. We approximate the equation (1.7) by

$$u_\lambda(t)+\int_0^t A(t-s)g_\lambda(s)u_\lambda(s)\,ds=f(t) \tag{3.1}$$

where $g_\lambda(s)$ is the Yosida approximation of $g(s)$ (when $s\in E$) (see [4]). But before we can treat this equation we must draw some conclusions from the hypothesis on $g(t)$:

LEMMA 1. *If* (2.14)–(2.16) *hold then* $Dg(t)$ *is independent of* t $(=Dg)$, $t\in E$ *and* $v\in L^p(0,T;H)\Rightarrow g_\lambda(t)v(t)\in L^p(0,T;H)$, $T>0$, $1\leqq p\leqq\infty$, $\lambda\in(0,\lambda_0]$.

Proof. Let $r,s\in E$ be such that $|r-s|\leqq T_2$. Then (2.15) and (2.16) imply that $|g_\lambda(r)x|\to\infty$ when $\lambda\to 0$ iff $|g_\lambda(s)x|\to\infty$ when $\lambda\to 0$ and this gives (see [4, p. 28]) $Dg(r)=Dg(s)$ which implies that $Dg(t)$ is independent of t. Let $T>0$, $p\in[1,\infty]$ and $\lambda\in(0,\lambda_0]$ be arbitrary and suppose $v\in L^p(0,T;H)$. It follows from (2.15) that e is continuous a.e. on $[0,T]$ and this together with (2.14) and (2.16) gives that for a given $x\in H$ $g_\lambda(t)x$ is a.e. continuous from the right in $[0,T]$. We conclude that

$g_\lambda(t)x$ is measurable. By (2.14) $g_\lambda(t)$: $H\to H$ is Lipschitz continuous (see [4, p. 28]) for fixed $t\in E$ and so $g_\lambda(t)v(t)$ is measurable because v is measurable by definition. From (2.16) and using the Lipschitz continuity of $g_\lambda(0)$ we deduce

$$|g_\lambda(t)v(t)|\leqq 2\max_{s\in[0,T]}|e(s)|L(|v(t)|+1)+\left(2\max_{s\in[0,T]}|e(s)|+1\right)$$
$$\cdot\left(|g_\lambda(0)x|+\frac{1}{\lambda}|v(t)|+\frac{1}{\lambda}|x|\right)$$

if $t\in E\cap[0,T]$ and $x\in H$. From this inequality the second conclusion follows and completes the proof of Lemma 1. □

Applying Lemma 1 and a standard argument made possible by the Lipschitz continuity of $g_\lambda(t)$: $H\to H$ that follows from (2.14) if $t\in E$, we deduce from (2.5) and (2.10) that there exists a function u_λ, absolutely continuous on every interval $[0,T]$ and such that (3.1) holds for every $t\geqq 0$. We also see that $G_\lambda(t)=g_\lambda(t)u_\lambda(t)\in L^\infty(0,T;H)$ for every $T>0$. Differentiating (3.1) yields

$$u_\lambda'(t)+A(0)G_\lambda(t)+\int_0^t A'(t-s)G_\lambda(s)\,ds=f'(t)\quad \text{a.e. } t\geqq 0. \tag{3.2}$$

From this equation we will under the hypothesis of the theorem extract some suitable bounds on G_λ and u_λ, which guarantee that we get a solution of (1.7) when we let $\lambda\to 0$. The solution will first be constructed on an interval $[0,T]$ where $T\leqq\min\{T_1,T_2\}$ and is chosen to satisfy the conditions (3.16), (3.18) and (3.29). We observe that T depends only on q, L, T_1, T_2 and the values of $A(t)$ on $[0,T_1]$.

The next lemma gives a crucial bound on G_λ.

LEMMA 2. *If* (2.5), (2.6), (2.10), (2.11), (2.13)–(2.17) *hold, then*

$$\sup_{\lambda\in(0,\lambda_0]}\|G_\lambda\|_{L^\infty(0,T)}<\infty. \tag{3.3}$$

Proof. Let $h>0$ and let $\lambda\in(0,\lambda_0]$ be arbitrary. By (2.13) and (3.2) we have

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}|u_\lambda(t+h)-u_\lambda(t)|^2&=(u_\lambda'(t+h)-u_\lambda'(t),\,u_\lambda(t+h)-u_\lambda(t))\\
&=-a_0(g_\lambda(t+h)u_\lambda(t+h)-g_\lambda(t+h)u_\lambda(t),\,u_\lambda(t+h)-u_\lambda(t))\\
&\quad-a_0(g_\lambda(t+h)u_\lambda(t)-g_\lambda(t)u_\lambda(t),\,u_\lambda(t+h)-u_\lambda(t))\\
&\quad+\Big(f'(t+h)-f'(t)-\int_0^{t+h}A'(t+h-s)G_\lambda(s)\,ds\\
&\quad+\int_0^t A'(t-s)G_\lambda(s)\,ds,\,u_\lambda(t+h)-u_\lambda(t)\Big),\quad \text{a.e. } t\in[0,T].\end{aligned}$$

When $t+h\in E$, $g_\lambda(t+h)$ is monotone and so by (2.13), (2.14) and the absolute continuity of u_λ integration of the preceding inequality over $(0, t)$ yields:

$$\tfrac{1}{2}|u_\lambda(t+h)-u_\lambda(t)|^2 \leqq \tfrac{1}{2}|u_\lambda(h)-u_\lambda(0)|^2 + \int_0^t \Big(a_0|g_\lambda(s+h)u_\lambda(s)-g_\lambda(s)u_\lambda(s)| + |f'(s+h)-f'(s)| + \Big|\int_0^{s+h} A'(s+h-r)G_\lambda(r)\,dr - \int_0^s A'(s-r)G_\lambda(r)\,dr\Big|\Big) \cdot |u_\lambda(s+h)-u_\lambda(s)|\,ds.$$

An application of Lemma A5 in [3] to the inequality above now gives

$$|u_\lambda(t+h)-u_\lambda(t)| \leqq |u_\lambda(h)-u_\lambda(0)| + a_0\int_0^t |g_\lambda(s+h)u_\lambda(s)-g_\lambda(s)u_\lambda(s)|\,ds + \int_0^t |f'(s-h)-f'(s)|\,ds + \int_0^t\int_0^s |A'(s+h-r)-A'(s-r)|\,|G_\lambda(r)|\,dr\,ds + \int_0^t \Big|\int_s^{s+h} A'(s+h-r)G_\lambda(r)\,dr\Big|\,ds, \qquad t\in[0,T]. \tag{3.4}$$

Our next purpose is to divide by h in (3.4) and let $h\to 0+$. To this end we need some estimates for the different terms on the right side of (3.4). Obtaining these estimates will occupy us until the relation (3.15). By (3.1)

$$\frac{1}{h}|u_\lambda(h)-u_\lambda(0)| \leqq \frac{1}{h}|f(h)-f(0)| + \frac{1}{h}\int_0^h |A(h-s)|\,|G_\lambda(s)|\,ds. \tag{3.5}$$

It follows from (2.6) that $\lim_{h\to 0+}|f'(h)|$ exists and so, using the absolute continuity of f,

$$\limsup_{h\to 0+}\frac{1}{h}|f(h)-f(0)| \leqq \limsup_{h\to 0+}\frac{1}{h}\int_0^h |f'(s)|\,ds < \infty. \tag{3.6}$$

By (2.7), (2.14), (2.15), (2.16) (with $x=f(0)$), (3.1) the Lipschitz continuity of $g_\lambda(s)\colon H\to H$ and the fact that $|g_\lambda(s)x| \leqq |g^0(s)x|$ if $x\in Dg(s)$ and $s\in E$ (see [3, p. 29]) we can deduce

$$\begin{aligned}|g_\lambda(s)u_\lambda(s)| &\leqq |g_\lambda(s)f(0)| + \frac{1}{\lambda}|u_\lambda(s)-f(0)| \\ &\leqq c_1|g_\lambda(0)f(0)| + c_2 + \frac{1}{\lambda}|u_\lambda(s)-u_\lambda(0)| \\ &\leqq c_3 + \frac{1}{\lambda}|u_\lambda(s)-u_\lambda(0)|, \qquad s\in E.\end{aligned}$$

This inequality combined with (3.5), (3.6) and the continuity of $|A(h-s)|$ and $u_\lambda(s)$ now yields

$$\limsup_{h\to 0+}\frac{1}{h}|u_\lambda(h)-u_\lambda(0)|\leqq c_4. \tag{3.7}$$

From (2.5), (2.10), (3.1) and an application of Hölder's inequality we obtain

$$\|u_\lambda\|_{L^\infty(0,T)}\leqq c_5+\|A\|_{L^1(0,T)}\|G_\lambda\|_{L^\infty(0,T)}, \tag{3.8}$$

and now using (2.16) we get

$$\begin{aligned}\int_0^t &|g_\lambda(s+h)u_\lambda(s)-g_\lambda(s)u_\lambda(s)|\,ds\\ &\leqq \int_0^t |e(s+h)-e(s)|[L(|u_\lambda(s)|+1)+|G_\lambda(s)|]\,ds\\ &\leqq (L\|A\|_{L^1(0,T)}\|G_\lambda\|_{L^\infty(0,T)}+c_6+\|G_\lambda\|_{L^\infty(0,T)})\\ &\qquad\cdot\int_0^{T-h}|e(s+h)-e(s)|\,ds,\qquad t\in[0,T-h].\end{aligned} \tag{3.9}$$

We have assumed of T that

$$0<T\leqq T_2 \tag{3.10}$$

holds and so (2.13), (2.15), (3.9) and an application of Lemma A1 in [3] yield

$$\begin{aligned}\limsup_{h\to 0+}\frac{1}{h}a_0\int_0^t &|g_\lambda(s+h)u_\lambda(s)-g_\lambda(s)u_\lambda(s)|\,ds\\ &\leqq (a_0Lq\|A\|_{L^1(0,T)}+a_0q)\|G_\lambda\|_{L^\infty(0,T)}+a_0qc_6,\qquad t\in[0,T).\end{aligned} \tag{3.11}$$

By (2.6) and another application of Lemma A1 in [3] we have

$$\limsup_{h\to 0+}\frac{1}{h}\int_0^t|f'(s+h)-f'(s)|\,ds\leqq \operatorname{var}(f;[0,T]),\qquad t\in[0,T). \tag{3.12}$$

The relation (2.11) and the same lemma that was used above imply

$$\begin{aligned}\limsup_{h\to 0+}\frac{1}{h}\int_0^t\int_0^s &|A'(s+h-r)-A'(s-r)||G_\lambda(r)|\,dr\,ds\\ &\leqq \|G_\lambda\|_{L^\infty(0,T)}\int_0^t \operatorname{var}(A';[0,T])\,ds\\ &\leqq \operatorname{var}(A';[0,T])T\|G_\lambda\|_{L^\infty(0,T)},\qquad t\in[0,T].\end{aligned} \tag{3.13}$$

We also have, adding and subtracting an additional term,

$$\begin{aligned}\frac{1}{h}\int_0^t\left|\int_s^{s+h}A'(s+h-r)G_\lambda(r)\,dr\right|\,ds &\leqq \int_0^t\left|\frac{1}{h}\int_s^{s+h}A'(s+h-r)G_\lambda(r)\,dr-A'(0+)G_\lambda(s)\right|\,ds\\ &\qquad+|A'(0+)|\int_0^t|G_\lambda(s)|\,ds.\end{aligned}$$

By (2.11) $\lim_{s\to 0+} A'(s) = A'(0+)$ exists and together with the integrability of G_λ this implies that the first integral on the right side in the preceding inequality converges to zero by the dominated convergence theorem. The second term can be estimated by using Hölder's inequality and so

$$\limsup_{h\to 0+} \frac{1}{h}\int_0^t \left|\int_s^{s+h} A'(s+h-r)G(r)\,dr\right| ds \leqq |A'(0+)|T\|G_\lambda\|_{L^\infty}, \qquad t\in[0,T). \tag{3.14}$$

Combining (3.4), (3.7), (3.11), (3.12), (3.13) and (3.14) we have

$$\begin{aligned}\|u_\lambda'\|_{L^\infty(0,T)} &\leqq \sup_{t\in[0,T)} \limsup_{h\to 0+} \frac{1}{h}|u_\lambda(t+h)-u_\lambda(t)| \\ &\leqq c_7 + (a_0 Lq\|A\|_{L^1(0,T)} + a_0 q + \operatorname{var}(A'[0,T])T \\ &\qquad + |A'(0+)|T)\|G_\lambda\|_{L^\infty}.\end{aligned} \tag{3.15}$$

We have assumed of T that

$$a_0 - \int_0^T |A'(s)|\,ds > 0 \tag{3.16}$$

holds. By (2.10) this is possible for $T>0$. Now we obtain from (2.6), (2.13) and (3.2)

$$\|G_\lambda\|_{L^\infty(0,T)} \leqq c_8 + \left(a_0 - \int_0^T |A'(s)|\,ds\right)^{-1} \|u_\lambda'\|_{L^\infty(0,T)}. \tag{3.17}$$

We have also assumed of T that it satisfies the condition

$$\begin{aligned} a_0 Lq\|A\|_{L^1(0,T)} + a_0 q + \operatorname{var}(A';[0,T])T + |A'(0+)|T \\ < a_0 - \int_0^T |A'(s)|\,ds.\end{aligned} \tag{3.18}$$

This is possible by (2.10), (2.11) and (2.15). The assertion of Lemma 2 now follows from (3.15), (3.17) and (3.18). □

It is now possible to deduce convergence of u_λ when $\lambda \to 0$. First we prove

LEMMA 3. *If* (2.10)–(2.14), (3.1) *and* (3.3) *hold then* $|\int_0^t [G_\lambda(s) - G_\mu(s)]\,ds| \to 0$ *when* $\lambda, \mu \to 0$ *and* $t\in[0,T]$.

Proof. If the assertion of Lemma 3 does not hold then we can conclude exactly as in [7] (see (3.16)–(3.24) in [7]) that there exist sequences $\{\lambda_n\}, \{\mu_n\} \to 0$, $d>0$, $t_0\in[0,T]$ such that if we put $h_n = G_{\lambda_n} - G_{\mu_n}$ we obtain

$$\left|\int_0^{t_0} h_n(s)\,ds\right| \geqq d, \qquad n=1,2,\cdots, \tag{3.19}$$

$$\left|\int_0^{t} h_n(s)\,ds\right| \leqq 2d, \qquad n=1,2,\cdots, \qquad t\in[0,t_0], \tag{3.20}$$

and

$$\left|\int_{s_1}^{s_2} h_n(s)\,ds\right| \leqq 4d, \qquad n=1,2,\cdots, \qquad s_1, s_2\in[0,t_0]. \tag{3.21}$$

We have assumed that T is such that

$$0 < T \leqq T_1 \tag{3.22}$$

holds. It follows by integration from (2.10) and (2.12) that $A(t)$ is symmetric if $t \in [0, T_1]$. This fact, (2.10), (2.12) and (3.22) yield for an arbitrary $z \in L^2(0, T; H)$ and $t \in [0, T]$

$$\begin{aligned}\int_0^t \left(z(s), \int_0^s A(s-r)z(r)\,dr\right) ds &= \frac{1}{2}\left(A(t)\int_0^t z(s)\,ds, \int_0^t z(s)\,ds\right)\\ &\quad -\frac{1}{2}\int_0^t \left(A'(s)\int_0^s z(r)\,dr, \int_0^s z(r)\,dr\right) ds\\ &\quad -\int_0^t\int_0^s \left(A'(s-r)\int_r^s z(p)\,dp, z(s)\right) dr\,ds.\end{aligned} \tag{3.23}$$

To see this, just perform an integration by parts two times and use the symmetry of A and A'.

It follows from (2.10)–(2.12) and (3.22) that we can construct a sequence $\{B_m\}$ with the following properties:

(3.24) $B_m \in W^{1,1}(0, T; L(H))$, $B_m \to A'$ pointwise a.e. in $[0, T]$ in the norm topology,

$$\int_0^T |B'_m(s)|\,ds \leqq \operatorname{var}(A'\colon [0, T]), \qquad m = 1, 2, \cdots, \tag{3.25}$$

and

$$B_m(s) \text{ is symmetric if } s \in [0, T], \qquad m = 1, 2, \cdots. \tag{3.26}$$

Now we claim that

$$\begin{aligned}&-\int_0^t\int_0^s \left(B_m(s-r)\int_r^s z(p)\,dp, z(s)\right) dr\,ds\\ &\quad = \frac{1}{2}\int_0^t\int_0^s \left(B'_m(r)\int_{s-r}^s z(p)\,dp, \int_{s-r}^s z(p)\,dp\right) dr\,ds\\ &\quad\quad -\frac{1}{2}\int_0^t \left(B_m(t-s)\int_s^t z(r)\,dr, \int_s^t z(r)\,dr\right) ds.\end{aligned} \tag{3.27}$$

This relation is a consequence of $B_m \in W^{1,1}(0, T; L(H))$, (3.26) an integration by parts and a change of variable.

Combining (3.3), (3.23) and (3.27) we obtain

$$\begin{aligned}&\int_0^t \left(h_n(s), \int_0^s A(s-r)h_n(r)\,dr\right) ds\\ &\quad = \frac{1}{2}\left(A(0)\int_0^t h_n(s)\,ds, \int_0^t h_n(s)\,ds\right)\\ &\quad\quad +\frac{1}{2}\left((A(t)-A(0))\int_0^t h_n(s)\,ds, \int_0^t h_n(s)\,ds\right)\\ &\quad\quad -\frac{1}{2}\int_0^t \left(A'(s)\int_0^s h_n(r)\,dr, \int_0^s h_n(r)\,dr\right) ds\\ &\quad\quad -\int_0^t\int_0^s \left((A'(s-r)-B_m(s-r))\int_r^s h_n(p)\,dp, h_n(s)\right) dr\,ds\end{aligned} \tag{3.28}$$

$$+\frac{1}{2}\int_0^t\int_0^s\left(B_m'(r)\int_{s-r}^s h_n(p)\,dp, \int_{s-r}^s h_n(p)\,dp\right)dr\,ds$$

$$+\frac{1}{2}\int_0^t\left((A'(t-s)-B_m(t-s))\int_s^t h_n(r)\,dr, \int_s^t h_n(r)\,dr\right)ds$$

$$-\frac{1}{2}\int_0^t\left(A'(t-s)\int_s^t h_n(r)\,dr, \int_s^t h_n(r)\,dr\right)ds.$$

We have assumed of T that it is such that

$$a_0-16s\,\mathrm{var}\,(A';[0,s])-24\int_0^s|A'(r)|\,dr\geqq b>0\quad\text{when } s\in[0,T] \tag{3.29}$$

holds. By (2.10) and (2.11) this is possible.

Using (2.10), (3.3), (3.19), (3.20), (3.21), (3.24), (3.25), (3.28), (3.29) and the dominated convergence theorem we conclude that it is possible to choose m such that for $n=1,2,\cdots$

$$\begin{aligned}&\int_0^{t_0}\left(h_n(s), \int_0^s A(s-r)h_n(r)\,dr\right)ds\\ &\qquad\geqq\frac{d^2}{2}a_0-2d^2\int_0^{t_0}|A'(s)|\,ds-8d^2t_0\,\mathrm{var}\,(A'\colon[0,T_0])\\ &\qquad\quad-10d^2\int_0^{t_0}|A'(s)|ds-\frac{bd^2}{4}\geqq\frac{bd^2}{4}>0.\end{aligned} \tag{3.30}$$

Write (3.1) with $\lambda=\lambda_n$, $\lambda=\mu_n$, take differences, multiply by h_n and integrate over $(0,t_0)$. This yields when using (3.30)

$$\begin{aligned}&\int_0^{t_0}(u_{\lambda_n}(s)-\mu_{\mu_n}(s), g_{\lambda_n}(s)u_{\lambda_n}(s)-g_{\mu_n}(s)u_{\mu_n}(s))\,ds\\ &\qquad\leqq-\frac{bd^2}{4}.\end{aligned} \tag{3.31}$$

But $g_\lambda(t)u_\lambda(t)\in g(t)J_\lambda(t)u_\lambda(t)$ a.e. $t\in[0,t_0]$ and so by the monotonicity of $g(t)$, implied by (2.14) when $t\in E$,

$$(-J_{\lambda_n}(t)u_{\lambda_n}(t)+J_{\mu_n}(t)u_{\mu_n}(t), g_{\lambda_n}(t)u_{\lambda_n}(t)-g_{\mu_n}(t)u_{\mu_n}(t))\leqq 0,\qquad\text{a.e. } t\in[0,t_0]. \tag{3.32}$$

Integrate (3.32) over $(0,t_0)$, add the result to (3.31) and use $u_\lambda(t)-J_\lambda(t)u_\lambda(t)=\lambda g_\lambda(t)u_\lambda(t)$ a.e. $t\in[0,t_0]$ and we get

$$\int_0^{t_0}(\lambda_nG_{\lambda_n}(s)-\mu_nG_{\mu_n}(s), G_{\lambda_n}(s)-G_{\mu_n}(s))\,ds\leqq-\frac{bd^2}{4}<0. \tag{3.33}$$

Recalling (3.3) and letting $\lambda_n,\mu_n\to 0$ we conclude that (3.33) is contradictory and so the assertion of Lemma 3 holds. □

Now we have the necessary tools to prove convergence of u_λ. Take $\lambda=\lambda_n,\mu_n$ for any $\{\lambda_n\},\{\mu_n\}\to 0$ in (3.1), take differences and perform an integration by parts justified by (2.10). This gives

$$u_{\lambda_n}(t)-u_{\mu_n}(t)=A(t)\int_0^t \Big(G_{\lambda_n}(s)-G_{\mu_n}(s)\Big)ds \\ -\int_0^t A'(t-s)\int_s^t \Big(G_{\lambda_n}(r)-G_{\mu_n}(r)\Big)\,dr\,ds. \tag{3.34}$$

It follows from Lemma 3, (2.10), (2.11), (3.3), (3.34) and an application of the dominated convergence theorem that

$$\lim_{n\to\infty} |u_{\lambda_n}(t)-u_{\mu_n}(t)|=0, \qquad t\in[0,T].$$

From this fact combined with (3.3), (3.8) and the dominated convergence theorem we can deduce that u_{λ_n} is a Cauchy sequence in $L^2(0,T;H)$ and so by the relation $J_\lambda(t)u_\lambda(t)=u_\lambda(t)-\lambda G_\lambda(t)$ and (3.3) it follows that there exists an $u\in L^2(0,T;H)$ such that

$$J_{\lambda_n}u_{\lambda_n}\to u \quad \text{in } L^2(0,T;H) \quad \text{when } n\to\infty. \tag{3.35}$$

From (3.3) we deduce that

$$\sup_{\lambda\in(0,\lambda]} \|G_\lambda\|_{L^2(0,T)}<\infty$$

and so by the weak sequential compactness of $L^2(0,T;H)$ there exist a subsequence of $\{\lambda_n\}$ (also denoted $\{\lambda_n\}$) and $w\in L^2(0,T;H)$ such that

$$G_{\lambda_n}\to w \quad \text{in } L^2(0,T;H)\text{(weakly)} \quad \text{when } n\to\infty. \tag{3.36}$$

In order to show that (1.8) and (1.9) hold, we need the following.

LEMMA 4. *Let* (2.14)–(2.16) *hold. Let* $D\hat{g}=\{u|u\in L^2(0,T;H),\ u(t)\in Dg$ *a.e. on* $[0,T]\}$ *and define* $\hat{g}$ *on* $D\hat{g}$ *by* $\hat{g}u=\{v|v\in L^2(0,T;H)v(t)\in g(t)u(t)$ *a.e. on* $[0,T]\}$. *Then* $\hat{g}$ *is a maximal monotone operator in* $L^2(0,T;H)$.

Proof. From the monotonicity of $g(t)$ a.e. follows that $\hat{g}$ is monotone in $L^2(0,T;H)$. To obtain the result that $\hat{g}$ is maximally monotone it is enough to show that $R(I+\lambda\hat{g})=L^2(0,T;H)$ for some $\lambda>0$. Let $\lambda\in(0,\lambda_0]$ and let $v\in L^2(0,T;H)$ be arbitrary. $g(t)$ is maximally monotone a.e. $t\in[0,T]$ and so $u(t)=J_\lambda(t)v(t)=(I+\lambda g(t))^{-1}v(t)$ is defined a.e. on $[0,T]$. It follows from Lemma 1 and the definition of $g_\lambda(t)$ that $u\in L^2(0,T;H)$ and the assertion of the lemma is obtained. □

As $\hat{g}$ is maximally monotone it is demiclosed so using (3.35) and (3.36) we have

$$u\in D\hat{g}, \qquad w\in\hat{g}u,$$

or by the definition of $\hat{g}$,

$$u(t)\in Dg, \qquad w(t)\in g(t)u(t) \quad \text{a.e. on } [0,T]. \tag{3.37}$$

Using (3.1) and the fact that the convolution operator with A as kernel by (2.10) is a bounded linear operator in $L^2(0,T;H)$ (and so also weakly continuous) we easily conclude from (3.35) and (3.36) that (3.38) holds a.e. on $[0,T]$.

This fact combined with (2.5) and (2.10) yields that u is absolutely continuous and so

$$u(t)+\int_0^t A(t-s)w(s)\,ds=f(t), \qquad t\in[0,T]. \tag{3.38}$$

By (3.37) we can without loss of generality assume that

$$u(T) \in Dg, \qquad T \in E. \tag{3.39}$$

As the set $\{v \in L^2(0, T; H) | \|v\|_{L^\infty(0,T)} \leqq C\}$ is closed in the strong topology of $L^2(0, T; H)$ and convex it is also closed in the weak topology and so (3.3) and (3.36) imply that

$$w \in L^\infty(0, T; H). \tag{3.40}$$

Now that we have proved the existence of a solution of (1.7) on the interval $[0, T]$ we use an induction argument to construct a solution on $[0, \infty)$. Suppose we have a solution of (1.7) on $[0, T_0]$ that also satisfies the conditions (3.38)–(3.40) for T replaced by T_0. Consider the equation

$$\begin{aligned} v(t) + \int_0^t A(t-s) g(T_0+s) v(s)\, ds &\ni f(t+T_0) \\ -\int_0^{T_0} A(t+T_0-s) w(s)\, ds &\overset{\text{def}}{=} f_0(t). \end{aligned} \tag{3.41}$$

Using (2.5), (2.6), (2.10), (2.11) and (3.40) we conclude that $f_0 \in W^{1,1}([0, T]; H)$ and that f_0' is of essentially bounded variation on $[0, T]$. (3.38) and (3.39) yield that $f_0(0) = u(T_0) \in Dg$. We can now apply the preceding arguments and construct a solution v of (3.41) on the interval $[0, T]$ such that (3.38)–(3.40) (with f replaced by f_0) are satisfied. If we define

$$\hat{u}(t) = \begin{cases} u(t) & \text{if } t \in [0, T_0] \\ v(t - T_0) & \text{if } t \in (T_0, T_0 + T] \end{cases}$$

then $\hat{u}$ is a solution of (1.7) on $[0, T_0 + T]$ such that (3.38)–(3.40) hold with T replaced by $T_0 + T$. The induction argument applies and gives a solution u on $[0, \infty)$ such that (3.38) and (3.40) hold for arbitrary T. The only thing left of the proof of the existence part of the theorem is to show that u is Lipschitz continuous on every interval $[0, T]$. By (2.5), (2.10) and (3.40) we can differentiate (3.38) which yields

$$u'(t) + A(0) w(t) + \int_0^t A'(t-s) w(s)\, ds = f'(t), \quad \text{a.e. } t \geqq 0.$$

Using (2.6), (2.10) and (3.40) it follows from this equation that $u' \in L^\infty(0, T; H)$, $\forall T > 0$ and so by the absolute continuity of u it follows that u is Lipschitz continuous on every interval $[0, T]$.

To prove uniqueness and thereby complete the proof we use the same method as in (3.74)–(3.83) in [7]. It is easily seen that this method can also be applied under the assumptions of this theorem. In particular note that no subdifferential properties of g are used.

4. Examples. Let Ω be a bounded open subset of R^n with sufficiently smooth boundary Γ and let $W^{m,p}$, $W_0^{m,p}$ stand for the Sobolev spaces. We shall consider a nonlinear differential operator of the form

$$Au = \sum_{|\alpha| \leqq m} (-1)^{|\alpha|} D^\alpha A_\alpha(x, u, \cdots, D^m u), \tag{4.1}$$

where $A_\alpha(x, z)$ are real functions defined on $\Omega \times R^N$. $A_\alpha(x, z)$ is measurable in x and continuous in z for all α. Assume there exist $p > \max\{1, (2n/(n+2m))\}$, $h \in L^q(\Omega)$,

$(1/p+1/q=1)$, and a constant C such that for all α

$$|A_\alpha(x,z)| \leqq C(|z|^{p-1}+h(x)), \quad \text{a.e. } x\in\Omega,\ z\in R^N. \tag{4.2}$$

For any $(z,y)\in R^N\times R^N$ and for almost every $x\in\Omega$ the following inequality is assumed to hold

$$\sum_{|\alpha|\leqq m}(A_\alpha(x,z)-A_\alpha(x,y))(z_\alpha-y_\alpha)\geqq 0. \tag{4.3}$$

Let $H=L^2(\Omega)$ and $V=W_0^{m,p}(\Omega)$. V' is the dual of V and $\langle\cdot,\cdot\rangle$ the duality pairing. If we identify H with its own dual we have $V\subset H\subset V'$ the imbeddings being continuous and dense. Define $\bar{A}$: $V\to V'$ by

$$\langle\bar{A}u,v\rangle=\sum_{|\alpha|\leqq m}\int_\Omega A_\alpha(x,u,\cdots,D^m u)D^\alpha v\,dx. \tag{4.4}$$

It follows from (4.2) and (4.3) that $\bar{A}$ is monotone and demicontinuous from V to V' (see [2, prop. 1.2, p. 49]). We assume that for some $\lambda>0$

$$\lim_{\|u\|_v\to\infty}\frac{\langle\bar{A}u,u\rangle+\lambda|u|_H^2}{\|u\|_v}=+\infty. \tag{4.5}$$

This is for instance the case if there exist $k\in L^1(\Omega)$ and positive constants c_2, c_3 such that

$$\sum_{|\alpha|\leqq m}A_\alpha(x,z)z_\alpha\geqq c_1\sum_{|\alpha|=m}|z_\alpha|^p-c_2|z_0|^2-k(x), \quad \text{a.e. } x\in\Omega. \tag{4.6}$$

If we define $gu=\bar{A}u$ on $Dg=\{u\in V|\bar{A}u\in H\}$ then we see just as in the proof of Theorem 2.5 p. 140 in [2] that g is a maximally monotone operator in H.

Suppose in addition that we have a function K: $[0,\infty)\to R$ that satisfies

(4.7) $\exists\varepsilon$ such that $K(t)\geqq\varepsilon>0$, $\forall t\geqq 0$.

(4.8) $\exists T_0>0$ such that var $(K;[T,T_0+T])\leqq q\varepsilon$, $q<1$, $\forall T\geqq 0$.

We define $g(t)=K(t)g$. Clearly $g(t)$ is maximally monotone in H for all $t\geqq 0$. We also have for any $\lambda>0$

$$\begin{aligned}|g_\lambda(t)x-g_\lambda(s)x|&=\frac{1}{\lambda}|J_\lambda(t)x-J_\lambda(s)x|\\&=\frac{1}{\lambda}|J_\lambda(s)(I+\lambda g(s))J_\lambda(t)x-J_\lambda(s)(I+\lambda g(t))J_\lambda(t)x|\\&\leqq|g(s)J_\lambda(t)x-g(t)J_\lambda(t)x|\leqq|K(s)-K(t)|\,|gJ_\lambda(t)x|\\&\leqq|K(s)-K(t)|\frac{1}{\varepsilon}|g_\lambda(t)x|.\end{aligned}$$

An application of Theorem 2 combined with the preceding remarks now yields

COROLLARY. *Assume* (2.1)–(2.3), (4.1)–(4.5), (4.7)–(4.8) *and*

(4.9) $f\in W^{1,1}(0,T;L^2(\Omega))$, $\forall T>0$,

(4.10) *f' is of essentially bounded variation on* $[0,T]$, $\forall T>0$, (4.11) *$f(0) \in L^2(\Omega)$ is such that the linear functional on $W_0^{m,p}(\Omega)$ defined by*

$$v \in W_0^{m,p}(\Omega) \to \sum_{|\alpha| \leqq m} \int_\Omega A_\alpha(x, f(0,x), \cdots, D^m f(0,x)) D^\alpha v(x)\, dx$$

belongs to $L^2(\Omega)$.

Then there exists a unique function $u: [0, \infty) \to L^2(\Omega)$ such that

(4.12) *u is Lipschitz continuous on every interval $[0, T]$,*

(4.13) *$u(t) \in W_0^{m,p}(\Omega)$ for a.e. $t \geqq 0$,*

(4.14) *there exists a function $w: [0, \infty) \to L^2(\Omega)$ such that $w \in L^\infty(0, T; L^2(\Omega))$, $\forall T > 0$,*

(4.15) *for* a.e. *$t \geqq 0$ and for all $v \in W_0^{m,p}(\Omega)$ we have*

$$\int_\Omega w(t,x) v(x)\, dx = \sum_{|\alpha| \leqq m} \int_\Omega K(t) A_\alpha(x, u(t,x), \cdots, D^m u(t,x)) D^\alpha v(x)\, dx$$

and

(4.16) *the following equation holds in $L^2(\Omega)$ for all $t \geqq 0$:*

$$u(t,x) + \int_0^t a(t-s) w(s,x)\, ds = f(t,x).$$

REFERENCES

[1] V. BARBU, *Nonlinear Volterra equations in Hilbert space*, this Journal, 6(1975), pp. 728–741.

[2] ———, *Nonlinear semigroups and differential equations in Banach spaces*, Editura Academiei, Bucaresti, Romania, and Noordhoff, Leyden, the Netherlands, 1976.

[3] ———, *On a nonlinear Volterra integral equation on a Hilbert space*, to appear in SIAM J. Math. Anal.

[4] H. BRÉZIS, *Operateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland, Elsevier, Amsterdam, 1973.

[5] M. G. CRANDALL AND A. PAZY, *Nonlinear evolution equations in Banach spaces*, Israel J. Math., 11(1972), pp. 57–94.

[6] L. C. EVANS, *Nonlinear evolution equations in an arbitrary Banach space*, Israel J. Math., 21 (1975), pp. 1–42.

[7] S-O. LONDEN, *On an integral equation in a Hilbert space*, to appear in SIAM J. Math. Anal.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0004 $01.00/0

MONOTONY PROPERTIES AND INEQUALITIES FOR GREEN'S FUNCTIONS FOR MULTIPOINT BOUNDARY VALUE PROBLEMS*

PHILIP HARTMAN†

Abstract. We derive monotony properties for $G(\cdot, s)$ and deduce inequalities for G, where $G(t, s)$ is the Green's function belonging to a nonsingular, disconjugate, linear nth order differential equation and to a multipoint boundary value problem.

1. Introduction. The object of this note is to derive monotony properties and inequalities for Green's functions $G(t, s)$ for N point boundary value problems associated with a real disconjugate nth order linear differential operator, i.e.,

$$Lx \equiv D^n x + \sum_{j=0}^{n-1} p_j(t) D^j x, \qquad D = D_t = d/dt, \tag{1.1}$$

$$D^{i-1}x(a_j) = 0 \quad \text{for } 1 \leqq j \leqq N, \quad 1 \leqq i \leqq m_j, \quad \sum_{j=1}^{N} m_j = n, \tag{1.2}$$

where $a = a_1 < \cdots < a_N = b$, $m_j > 0$ is an integer, and $N \geqq 2$. Condition (1.2) means that $x(t)$ has a zero of order at least m_j at $t = a_j$. In this paper, unless otherwise stated, we make the following hypothesis.

(H) $p_j(t) \in C^0[a, b]$, *for* $0 \leqq j < n$, *is real-valued and* (1.1) *is disconjugate on* $[a, b]$.

The operator (1.1) is called *disconjugate* on $[a, b]$ if no solution $x(t) \not\equiv 0$ of $Lx = 0$ has n zeros on $[a, b]$. In this definition of disconjugacy, it is immaterial whether or not zeros are counted with their mutliplicities when the coefficients p_j are real-valued; for a more general result, see Hartman [5] and, for a simple proof for the case at hand, see Opial [8].

Our inequalities for the Green's function $G(t, s)$ implied by Corollary 2.1 and Proposition 2.1 below reduce, for the case $N = 2$, to those of Bates and Gustafson [1]. Our proof is simpler than that of [1] and is obtained with a minimum of calculation. Calculation is minimized by a systematic use of the adjoint boundary value problem and the avoidance of particular fundamental sets of solution of L and L^* in § 5. Our proof is particularly simple in the case $N = 2$ where the adjoint boundary value problem is obtained by replacing L by its formal adjoint L^* and m_1, m_2 by $n - m_1$, $n - m_2$.

For applications, it is useful to know bounds for the derivatives $D_t^k G(t, s)$ of the Green's function, but these can be obtained from bounds for G and bounds for the coefficients of L; cf., e.g., [4, pp. 480–481].

Some monotony properties of $G(t, s)$ are given in Theorem 2.1_k. The case $N = 2$ and $k = 1$ is implicit in [1]; see the remarks below following Theorem 2.1_1. Other monotony properties are given in Theorem 2.2.

The referee has called my attention to the paper [2] which has appeared since this one was submitted for publication. In it, Bates and Gustafson show, for example, that $|G(t, s)| \leqq \max \{|G_k(t, s)| : 0 < k < n\}$, where $G_k(t, s)$ is the Green's function for the two point boundary value problem belonging to $N = 2$, $m_1 = k$, $m_2 = n - k$ in (1.2).

---

* Received by the editors September 20, 1976, and in revised form January 24, 1977.

† Mathematics Department, Johns Hopkins University, Baltimore, Maryland 21218. This study was supported by the National Science Foundation under Grant MPS75-15733.

2. Statement of results. Recall that if (1.1) is disconjugate on $[a, b]$, then there exists a (unique) Green's function $G(t, s) \in C^0([a, b] \times [a, b])$ with the property that if $f(t) \in C^0[a, b]$, then

$$x(t) = \int_a^b G(t, s) f(s)\, ds$$

is the (unique) solution of $Lx = f$ satisfying the boundary conditions (1.2). The function G is determined by the conditions: For fixed $s \in (a, b)$, $x(t) = G(t, s)$ satisfies $Lx = 0$ on $a \leqq t < s$ and $s < t \leqq b$, the boundary conditions (1.2), and $D^{i-1}x(s+0) - D^{i-1}x(s-0) = 0$ or 1 according as $1 \leqq i < n$ or $i = n$.

THEOREM 2.1_1. *Assume hypothesis* (H). *Let* $1 \leqq J < N$ *and* $x = x_{J1}(t)$ *be the solution of* $Lx = 0$ *determined by the conditions*

$$D^{i-1}x(a_j) = 0 \quad \text{for } j \neq J \quad \text{and} \quad 1 \leqq i \leqq m_j, \tag{2.1}$$

$$D^{i-1}x(a_J) = 0 \text{ or } 1 \quad \text{according as} \quad 1 \leqq i < m_J \text{ or } i = m_J. \tag{2.2}$$

Let $s \in (a, b)$ *be fixed. Then*

$$(-1)^{\sigma(J)} D_t\{G(t, s)/x_{J1}(t)\} > 0 \quad \text{for } a \leqq t \leqq b, \quad \sigma(J) = \sum_{j=J+1}^{N} m_j, \tag{2.3}$$

holds unless either $J = N$, $m_N = 1$, $a_{N-1} < s < b$, *and*

$$(-1)^{\sigma(J)} D_t\{G/x_{J1}\} \equiv 0 \text{ or } > 0 \quad \text{according as} \quad a \leqq t \leqq s \text{ or } s < t \leqq b, \tag{2.4}$$

or $J = 1$, $m_1 = 1$, $a < s < a_2$, *and*

$$(-1)^{\sigma(J)} D_t\{G/x_{J1}\} > 0 \text{ or } \equiv 0 \quad \text{according as} \quad a \leqq t < s \text{ or } s \leqq t \leqq b. \tag{2.5}$$

For $N = 2$, Bates and Gustafson [1] essentially give Theorem 2.1_1 (cf. the proof of their Lemma 3.4, pp. 333–334) but, because of a misstatement of an implication of Coppel [3, p. 108] they seem to assert that (2.3) always holds.

Since $G(t, s)/x_{J1}(t) \to 0$ as $t \to a_J$ and $G(t, s)/x_{J1}(t) \to D_t^{m_j} G(a_j, s)/D^{m_j} x_{J1}(a_j)$ as $t \to a_j (\neq a_J)$, Theorem 2.1_1 implies

COROLLARY 2.1. *For* $j = 1, \cdots, N$, *put*

$$y_j(s) = D_t^{m_j} G(a_j, s) \quad \text{for } a \leqq s \leqq b. \tag{2.6}$$

Then, for fixed $s \in (a, b)$,

$$0 \leqq (-1)^{\sigma(J)} G/x_{J1} \leqq (-1)^{\sigma(J)} y_j(s)/D^{m_j} x_{J1}(a_j) \quad \text{for } a_J \leqq t \leqq a_j, \tag{2.7}$$

$$(-1)^{\sigma(J)} G/x_{J1} \geqq (-1)^{\sigma(J)} y_j(s)/D^{m_j} x_{J1}(a_j) \geqq 0 \quad \text{for } a_J \leqq a_j \leqq t \leqq b, \tag{2.8}$$

$$0 \geqq (-1)^{\sigma(J)} G/x_{J1} \geqq (-1)^{\sigma(J)} y_j(s)/D^{m_j} x_{J1}(a_j) \quad \text{for } a_j \leqq t \leqq a_J, \tag{2.9}$$

$$(-1)^{\sigma(J)} G/x_{J1} \leqq (-1)^{\sigma(J)} y_j(s)/D^{m_j} x_{J1}(a_j) \leqq 0 \quad \text{for } a \leqq t \leqq a_j \leqq a_J. \tag{2.10}$$

Note that $y_j(s) \in C^0[a, b]$. This is clear if $m_j < n-1$. Also if $m_j = n-1$, then $N = 2$ and a_j is a or b, and the continuity of $y_j(s)$ is again clear.

This corollary will be much more useful for giving bounds for G if we give a more "explicit" characterization of the function y_j in (2.6):

PROPOSITION 2.1. *Assume that* $p_j(t) \in C^j[a, b]$ *for* $j = 0, \cdots, n-1$ *and define the operator* L^* *by*

$$L^* x^* \equiv (-1)^n D^n x^* + \sum_{j=0}^{n-1} (-1)^j D^j [p_j(t) x^*]. \tag{2.11}$$

Let $X^*(t, s)$ *be the Cauchy function for* (2.11), *i.e.,* $x^*(t)=X^*(t, s)$ *is the solution of* $L^*x^*=0$ *satisfying the initial conditions*

$$(2.12)\qquad D^{i-1}x^*(s)=0 \text{ or } 1 \quad \text{according as} \quad 1\leqq i<n \text{ or } i=n.$$

Then there exists a function $y_j^0(t)$ *with the properties that* $x^*(t)=y_j^0(t)$ *is a solution of* $L^*x^*=0$ *on* (a_k, a_{k+1}) *for* $k=1,\cdots,N-1$ *and satisfies the boundary conditions*

$$(2.13_1)\qquad D^{i-1}x^*(a_1)=0 \qquad \text{for } 1\leqq i\leqq n-m_1,$$

$$(2.13_k)\qquad D^{i-1}x^*(a_k+0)=D^{i-1}x^*(a_k-0) \quad \text{for } 1<k<N, \quad 1\leqq i\leqq n-m_k,$$

$$(2.13_N)\qquad D^{i-1}x^*(a_N)=-D_s^{m_j}D_t^{i-1}X^*(b, a_j) \quad \text{for } 1\leqq i\leqq n-m_N;$$

and there exists a function $y_j^1(t)$ *with the properties that* $x^*(t)=y_j^1(t)$ *is a solution of* $L^*x^*=0$ *on* (a_k, a_{k+1}) *for* $k=1,\cdots,N-1$ *and satisfies the boundary conditions*

$$(2.14_1)\qquad D^{i-1}x^*(a_1)=D_s^{m_j}D_t^{i-1}X^*(a, a_j) \quad \text{for } 1\leqq i\leqq n-m_1,$$

$$(2.14_k)\qquad D^{i-1}x^*(a_k+0)=D^{i-1}x^*(a_k-0) \quad \text{for } 1<k<N, \quad 1\leqq i\leqq n-m_k,$$

$$(2.14_N)\qquad D^{i-1}x^*(a_N)=0 \qquad \text{for } 1\leqq i\leqq n-m_N.$$

Furthermore, the function (2.6) *is given by*

$$(2.15)\qquad y_j(t)=y_j^0(t) \quad \text{for } a\leqq t<a_j \quad \text{and} \quad y_j(t)=y_j^1(t) \quad \text{for } a_j<t\leqq b.$$

Remark. For the case $j=N$, we have that $x^*(t)=y_N(t)=y_N^0(t)$ for $a\leqq t\leqq b$ satisfies (2.13_k) for $1\leqq k<N$, while (2.13_N) becomes

$$(2.13'_N)\qquad D^{i-1}x^*(a_N)=0 \text{ or } (-1)^{m_N} \quad \text{according as } 1\leqq i<n-m_N \text{ or } i=m_N,$$

Correspondingly for $j=1$, $x^*(t)=y_1(t)=y_1^1(t)$ for $a\leqq t\leqq b$ satisfies

$$(2.14'_1)\qquad D^{i-1}x^*(a_1)=0 \text{ or } (-1)^{m_1} \quad \text{according as} \quad 1\leqq i<n-m_1 \text{ or } i=n-m_1,$$

and (2.14_k) for $1<k\leqq N$. The reduction of (2.13_N) to $(2.13'_N)$ and (2.14_1) to $(2.14'_1)$ follows from

$$(2.16)\qquad D_s^{k-1}D_t^{i-1}X^*(s, s)=0 \text{ or } (-1)^{k-1} \quad \text{according as} \quad i+k<n \text{ or } i+k=n;$$

cf; e.g., [1, p. 329].

Remark. The assumption $p_j\in C^j[a, b]$ (instead of $p_j\in C^0[a, b]$) is for convenience only and can be avoided if Proposition 2.1 is suitably reformulated.

If $m_J>1$, then Theorem 2.1_1 is the first of a sequence $(k=1,\cdots,m_J)$ of statements concerning "monotony" properties of $G(\cdot, s)$. For if $W(u_1,\cdots,u_i)$ denotes the Wronskian determinant of the i functions $u_1,\cdots,u_i$, then $D_t\{G/x_{J1}\}=x_{J1}^{-2}W(x_{J1}, G(\cdot, s))$.

THEOREM 2.1_k. *Assume hypothesis* (H). *Let* $1\leqq J\leqq N$, $1\leqq k\leqq m_J$, *and* $x=x_{Jk}(t)$ *be a solution of* $Lx=0$ *determined by conditions*

$$(2.17)\qquad D^{i-1}x(a_j)\neq 0 \quad \text{for } j\neq J, \quad 1\leqq i\leqq m_j,$$

$$(2.18)\qquad D^{i-1}x(a_J)=0 \text{ or } 1 \quad \text{according as} \quad 1\leqq i\leqq m_J-k \text{ or } i=m_J-k+1,$$

while $D^{i-1}x(a_J)$ *is arbitrary for* $m_J-k+1<i\leqq m_J$ *if* $k>1$. (*In particular,* x_{Jk} *has a zero of order exactly* m_J-k *at* $t=a_J$.) *Let* $s\in(a, b)$ *be fixed and*

$$(2.19)\qquad W_J^k(t, s)=W(x_{Jk}, x_{J,k-1},\cdots,x_{J1}, G(t, s)).$$

Then the inequality

$$(2.20)\qquad (-1)^{\sigma(J)}x_{Jk}^{-k-1}(t)W_J^k(t, s)>0 \quad \text{for } a\leqq t\leqq b$$

holds unless either $J=N$, $m_J=k$, $a_{N-1}<s<b$, *and*

(2.21) $$(-1)^{\sigma(J)}x_{Jk}^{-k-1}W_J^k\equiv 0 \text{ or } >0 \quad \text{according as} \quad a\leqq t\leqq s \text{ or } s<t\leqq b,$$

or $J=1$, $m_J=k$, $a<s<a_2$, *and*

(2.22) $$(-1)^{\sigma(J)}x_{Jk}^{-k-1}W_J^k>0 \text{ or } \equiv 0 \quad \text{according as} \quad a\leqq t<s \text{ or } s\leqq t\leqq b.$$

The proof of Theorem 2.1_k, and hence of Corollary 2.1, give the following extensions of these results.

COROLLARY 2.2. *Assume the analogue of hypothesis* (H) *in which* $[a,b]$ *is replaced by* (a,b) [*or* $(a,b]$ *or* $[a,b)$] *and let* $a<a_1<\cdots<a_N<b$ [*or* $a<a_1<\cdots<a_N\leqq b$ *or* $a\leqq a_1<\cdots<a_n<b$]. *For* $a_1<s<a_N$, *extend* $x(t)=G(t,s)$ *to* (a,b) [*or* $(a,b]$ *or* $[a,b)$] *as a solution of* $Lx=0$ *without discontinuities of its derivatives at* $t=a_1$, a_N. *Then, for* $s\in(a_1,a_N)$, *instead of* $s\in(a,b)$, *Theorem* 2.1_k *and Corollary* 2.1 *are valid for* $t\in(a,b)$ [*or* $(a,b]$ *or* $[a,b)$].

More monotony properties are contained in the following generalization of Theorem 2.1_k.

THEOREM 2.2. *Assume hypothesis* (H). *Let* $1\leqq J(1)<J(2)<\cdots<J(I)\leqq N$ *and* $1\leqq k(i)\leqq m_{J(i)}$ *for* $1\leqq i\leqq I$. *Let* $x=x_{J(i)k}(t)$ *be a solution of* $Lx=0$ *defined as in Theorem* 2.1_k *with* $(J,k)=(J(i),k)$, $1\leqq k\leqq k(i)$. *Put*

(2.23) $$W_{J(1)\cdots J(I)}^{k(1)\cdots k(I)}(t,s)=W(x_{J(1)k(1)},x_{J(1),k(1)-1},\cdots,x_{J(1)1},x_{J(2)k(2)},\cdots,x_{J(I)1},G),$$

so that (2.23) *is the Wronskian of* $K+1$ *functions* (*of* t), *where*

(2.24) $$K=k(1)+\cdots+k(I).$$

Define $S(t,s)$ *by*

(2.25) $$S(t,s)=\left\{\prod_{i=1}^{I}(t-a_{J(i)})^{m_{J(i)}-k(i)}\prod_{j\neq J(i)}(t-a_j)^{m_j}\right\}^{-K-1}W_{J(1)\cdots J(I)}^{k(1)\cdots k(I)}(t,s).$$

Then S is continuous, and

(2.26) $$(-1)^{\tau}S>0 \quad \text{for } a\leqq t\leqq b,\quad \tau=\sum_{i=1}^{I}k(i)\left[\sigma(J(i))+\sum_{j=i+1}^{I}k(j)\right],$$

unless either $(J(1),\cdots,J(I))=(N-I+1,\cdots,N)$, $a_{N-I}<s<b$, $m_{J(i)}=k(i)$, *and*

(2.27) $$S\equiv 0 \text{ or } (-1)^{\tau}S>0 \quad \text{according as} \quad a\leqq t\leqq s \text{ or } s<t\leqq b$$

or $(J(1),\cdots,J(I))=(1,\cdots,I)$, $a<s<a_{I+1}$, $m_{J(i)}=k(i)$, *and*

(2.28) $$(-1)^{\tau}S>0 \text{ or } S\equiv 0 \quad \text{according as} \quad a\leqq t<s \text{ or } s\leqq t\leqq b.$$

Remark. We can also extend this result as Theorem 2.1_k is extended in Corollary 2.2.

Theorem 2.1_k, $1\leqq k\leqq m_J$, will be proved in § 3, Theorem 2.2 in § 4, and Proposition 2.1 in § 5.

3. On Theorem 2.1_k, $1\leqq k\leqq m_J$. The proof will depend on the following lemma.

LEMMA 3.1. *Assume hypothesis* (H). *Let* $a<s<b$ *and let* $x=z(t)$ *be a solution of* $Lx=0$ *on* $[a,s)$ *and* $(s,b]$ *satisfying* $D^{i-1}z(s+0)=D^{i-1}z(s-0)$ *for* $1\leqq i<n$. *If* $z(t)$ *has at least* $n+1$ *zeros* (*counting multiplicities*), *then either* $z(t)\equiv 0$ *on* $[a,s]$ *or on* $[s,b]$.

This is implied by the proofs of Lemma 16 and Theorem 11 in Coppel [3, p. 108].

Proof of Theorem 2.1_k. First, note that if $m_j=n-1$ for some j, then $N=2$ and

$a_1 = a$ and $a_2 = b$, so that there is no difficulty about the existence and continuity of $D_t^{m_j}G(t, s)$ at $t = a_j$. From the fact that

$$G(t, s)/\prod_{j=1}^{N} (t-a_j)^{m_j} > 0 \quad \text{for } a \leqq t \leqq b, \quad a < s < b$$

(Levin [7] and Pokornyĭ [9]; cf. Coppel [3, p. 108]), it follows that

$$y_J(s)/(m_J!) \prod_{j \neq J} (a_J - a_j)^{m_j} > 0, \quad \text{where } y_J(s) = D_t^{m_J}G(a_J, s).$$

Hence

$$(3.1) \qquad (-1)^{\sigma(J)} y_J(s) > 0 \quad \text{for } a < s < b.$$

As $t \to a_J$, we also have, for $0 \leqq i \leqq m_J$,

$$(3.2) \qquad D_t^i G(t, s) \sim y_J(s)(t-a_J)^{m_J - i}/(m_J - i)!;$$

for $0 \leqq i \leqq k$,

$$(3.3) \qquad D_t^i\{G(t, s)/x_{Jk}(t)\} \sim y_J(s)k!(t-a_J)^{k-i}/(k-i)!;$$

and for $1 \leqq j \leqq k$, $0 \leqq i \leqq k-j$,

$$(3.4) \qquad D^i\{x_{Jj}/x_{Jk}\} \sim (k-j)!(t-a_J)^{k-j-i}/(k-j-i)!.$$

A standard identity for Wronskians gives

$$(3.5) \qquad x_{Jk}^{-k-1} W_J^k = W(1, x_{J,k-1}/x_{Jk}, \cdots, x_{J1}/x_{Jk}, G/x_{Jk});$$

cf., e.g., [6, p. 310]. By (3.3) and (3.4), the right side of (3.5) is the determinant of a matrix which at $t = a_J$ has zero entries above the main diagonal, and diagonal elements $1, 1!, 2!, \cdots, (k-1)!, y_J(s)k!$. Thus, by (3.1),

$$(3.6) \qquad (-1)^{\sigma(J)} x_{Jk}^{-k-1}(a_J) W_J^k(a_J, s) > 0 \quad \text{for } a < s < b.$$

Suppose that, for some fixed $s \in (a, b)$, there exists a t_0 such that $x_{Jk}^{-k-1} W_J^k = 0$ at $t = t_0$. Then $t_0 \neq a_J$ by (3.6). If $t_0 \neq a_1, \cdots, a_N$, then there exist constants $\alpha_1, \cdots, \alpha_k$ and β (not all 0) such that

$$(3.7) \qquad z(t) = \sum_{j=1}^{k} \alpha_j x_{Jj}(t) - \beta G(t, s)$$

has a zero of order $k+1$ at $t = t_0$, a zero of order m_j at $t = a_j (\neq a_J)$, and a zero of order $m_J - k$ at $t = a_J$. If $t_0 = a_i (\neq a_J)$, then constants $\alpha_1, \cdots, \alpha_k, \beta$ (not all 0) can be chosen so that z has a zero of order $m_i + k$ at $t = a_i$ (cf. (3.5)), a zero of order m_j at $t = a_j (\neq a_i, a_J)$, and a zero of order $m_J - k$ at $t = a_J$. In either case, z has at least $n+1$ zeros on $[a, b]$. Hence $z \equiv 0$ on either $[a, s]$ or $[s, b]$, by Lemma 3.1.

Note that $\beta \neq 0$, for otherwise $z = \alpha_1 x_{J1} + \cdots + \alpha_k x_{Jk} \equiv 0$ on $[a, b]$. In this case, $\beta = \alpha_1 = \cdots = \alpha_k = 0$ since $x_{J1}, \cdots, x_{Jk}$ are linearly independent. But this contradicts $(\alpha_1, \cdots, \alpha_k, \beta) \neq 0$. Also, we have $(\alpha_1, \cdots, \alpha_k) \neq 0$.

Suppose, for example, that $z \equiv 0$ on $[a, s]$. Then $z \neq 0$ on $(s, b]$, for otherwise $G = \sum \alpha_j x_{Jj}/\beta$ on $a \leqq t \leqq b$ does not satisfy the boundary conditions (1.2), in particular, the condition at $t = a_J$. Consider, for a moment, only $t \in [s, b]$. Then z has a zero of exactly order $n-1$ at $t = s$ (in view of $G = \sum \alpha_j x_{Jj}/\beta$ for $a \leqq t < s$ and the jump conditions on $D^{i-1}G$ at $t = s$). Since $z \not\equiv 0$ on $[s, b]$, it follows from the disconjugacy of (1.1) that $z(t) \neq 0$ on $(s, b]$.

Consequently, $J=N$, $m_N=k$ and $a_{N-1}<s$. For if $J\neq N$, then z has a zero of order $m_N\geqq 1$ at $t=a_N(=b)$, and if $J=N$ and $m_N>k$, then z has a zero of order $m_N-k\geqq 1$ at $t=a_N(=a_J=b)$.

Suppose, if possible, that there exists a $t_1\in(s,b]$, such that $x_{Jk}^{-k-1}W_J^k=0$ at t_1. In particular, $t_1\neq a_J(=b)$ by (3.6). Then there exist constants $\gamma_1,\cdots,\gamma_k$ and δ (not all 0) such that $z_1=\sum\gamma_j x_{Jj}-\delta G$ has a zero of at least order $k+1$ at t_1. As above, $\delta\neq 0$ and $(\gamma_1,\cdots,\gamma_k)\neq 0$. For convenience, suppose that $\beta=\delta=1$.

Now consider $t\in[a,b]$. On $[a,s)$, $z_1\not\equiv 0$, for otherwise $\beta=\delta=1$ implies that $z\equiv z_1$ vanishes at $t_1\in(s,b)$. Also, $z_1\not\equiv 0$ on $(s,b]$, for otherwise $G=\sum\alpha_j x_{Jj}$ on $[a,s)$ and $G=\sum\alpha_j x_{Jj}$ on $(s,b]$ does not satisfy the boundary conditions (1.2), in particular, the condition at $a_J(=b)$. Since z_1 has at least $(n+1)$ zeros on $[a,b]$, we have a contradiction (Lemma 3.1). This proves Theorem 2.1_k.

4. Proof of Theorem 2.2. We first show that if Theorem 2.2 is modified by replacing ">0" by "≠0" in (2.26)–(2.28), then the proof of the modified statement can be obtained by the arguments of § 3. To this end, we verify that S is continuous and not 0 at $t=a_{J(1)}$. It is sufficient to prove that

$$x_{J(1)k(1)}^{-K-1}W_{J(1)\cdots J(I)}^{k(1)\cdots k(I)}=W(1,x_{J(1),k(1)-1}/x_{J(1)k(1)},\cdots,G/x_{J(1)k(1)}) \tag{4.1}$$

is continuous and not 0 at $t=a_{J(1)}$. The continuity is clear. If it vanishes at $t=a_{J(1)}$, then there exist constants $\alpha_{J(i)k}$ and β, not all 0, such that $z(t)/x_{J(1)k(1)}$, where

$$z(t)=\sum_{i=1}^{I}\sum_{k=1}^{k(i)}\alpha_{J(i)k}x_{J(i)k}(t)-\beta G(t,s),$$

has a $(K+1)$-fold zero at $t=a_{J(1)}$. Thus $z(t)$ has a zero of order $K+m_{J(1)}-k(1)$ at $t=a_{J(1)}$, a zero of order $m_{J(i)}-k(i)$ at $t=a_{J(i)}$, $i>1$, and a zero of order m_j at $t=a_j$, $j\neq J(i)$. Thus $z(t)$ has $n+1$ zeros on $[a,b]$ and, by Lemma 3.1, $z(t)\equiv 0$ on $[a,s)$ or $[s,b]$.

Arguing as in § 3, we see that $\beta\neq 0$. If $z\equiv 0$ on $[a,s]$, then $a_{J(i)}\geqq s$ for otherwise $G(\cdot,s)$ does not satisfy the boundary condition (1.2) at $a_{J(i)}$. Also $z(t)\neq 0$ on $(s,b]$, with a zero of order $n-1$ at s, so that $a_{J(i)}>s$ and $k(i)=m_{J(i)}$, $1\leqq i\leqq I$, and $a_j<s$ if $j\neq J(i)$. In particular, (5.1) does not vanish at $t=a_{J(1)}$.

Analogous arguments show that S is continuous and not 0 at $t=a_{J(i)}$, $1\leqq i\leqq I$. The proof of the modified Theorem 2.2 now follows that of Theorem 2.1_k.

We now proceed to show that $\operatorname{sgn} S=(-1)^\tau$ when $S\neq 0$. On the t-intervals on which $S\neq 0$ in (2.26), (2.27) or (2.28), sgn S does not depend on the choice of the solutions $x_{J(i)k}$, $1<k\leqq k(i)$. For if $\{x_{J(i)k}\}$, $\{\tilde{x}_{J(i)k}\}$ are two admissible sets of solutions, then $\{\tau x_{J(i)k}+(1-\tau)\tilde{x}_{J(i)k}\}$ is also (for a fixed τ, $0\leqq\tau\leqq 1$). But the corresponding sgn S does not depend on τ (by continuity considerations). Also sgn S does not depend on $s\in(a,b)$ or the operator L, and can be determined by considering the trivial operator $L=D^n$. In fact, if one considers the set of linear differential operators $L=D^n+\sum p_j(t)D^j$ as a metric space with dist $(L,\tilde{L})=\sum\{\max|p_j-\tilde{p}_j|:a\leqq t\leqq b\}$, then the subset of disconjugate operators is open and connected, and contains $L=D^n$; cf. Coppel [3, pp. 95 and 107]. Thus sgn S is given by

$$\operatorname{sgn} V(t),\quad \text{where } V(t)=H(t)^{-K-1}W(y_{J(1)k(1)},\cdots,y_{J(I)1},P), \tag{4.2}$$

$$H(t)=\prod_{i=1}^{I}(t-a_{J(i)})^{m_{J(i)}-k(i)}\prod_{j\neq J(i)}(t-a_j)^{m_j}, \tag{4.3}$$

and $y_{Jk}(t)$, $P(t)$ are the polynomials

$$y_{Jk}(t)=(-1)^{\sigma(J)}(t-a_J)^{m_J-k}\prod_{j\neq J}(t-a_j)^{m_j}/(m_J-k)!,\qquad P(t)=\prod_{j\neq 1}^{N}(t-a_j)^{m_j}/n!.$$

For if $\tilde{G}$, $\tilde{W}_{J(1)\cdots J(I)}^{k(1)\cdots k(I)}$ belong to $\tilde{L}=D^n$, as G, $W_{J(1)\cdots J(I)}^{k(1)\cdots k(I)}$ belong to L, then $y=P(t)$ is the solution of $D^n y=1$ and (1.2), so that

$$P(t)=\int_a^b \tilde{G}(t,s)\,ds$$

and

$$W(y_{J(1)k(1)},\cdots,y_{J(I)1},P)=\int_a^b \tilde{W}_{J(1)\cdots J(I)}^{k(1)\cdots k(I)}(t,s)\,ds.$$

Consequently, it only remains to show that (4.2) is $(-1)^\tau$.

Note that if $z_{J(i)k}=y_{J(i)k}/H$ and $Q=P/H$, i.e.,

$$z_{J(i)k}=(-1)^{\sigma(J(i))}(t-a_{J(i)})^{k(i)-k}\prod_{j\neq i}(t-a_{J(j)})^{k(j)}/(m_{J(i)}-k)!,$$

$$Q=\prod_{i=1}^{I}(t-a_{J(i)})^{k(i)}/n!,$$

then a standard identity for Wronskians (cf. [6, p. 310]) gives the first of the relations

$$V(t)=W(z_{J(1)k(1)},\cdots,z_{J(I)1},Q)=(K!/n!)W(z_{J(1)k(1)},\cdots,z_{J(I)1}).$$

The last relation follows from the fact that the last row of the $(K+1)\times(K+1)$ matrix with determinant $W(z_{J(1)k(1)},\cdots,z_{J(I)1},Q)$ is $(0,\cdots,0,K!/n!)$, for Q is a polynomial of degree K with leading coefficient $1/n!$ and $z_{J(i)k}$ is a polynomial of degree less than K.

It is clear that $W_1\equiv W(z_{J(1)k(1)},\cdots,z_{J(I)1})\neq 0$ since it is the Wronskian of a fundamental set of solutions of $D^K z=0$. At $t=a_{J(1)}$, the kth diagonal element for $1\leqq k\leqq k(1)$ is

$$(-1)^{\sigma(J(1))}(k(1)-k)!\prod_{j\neq 1}(a_{J(1)}-a_{J(j)})^{k(j)}/(m_{J(1)}-k)!=(-1)^{\tau(1)}\rho,$$

where $\rho>0$ is a positive number and

$$\tau(i)=\sigma(J(i))+\sum_{j=i+1}^{I}k(j).$$

Also, at $t=a_{J(1)}$, the elements above and to the right of these diagonal elements vanish. Furthermore, the elements in the $(K-k(1))\times(K-k(1))$ matrix in the lower right corner are precisely $k(1)!$ times the elements in the matrix of the determinant $W_2\equiv W(z^1_{J(2)k(2)},\cdots,z^1_{J(I)1})$ at $t=a_{J(1)}$, where $z^1_{J(i)k}=z_{J(i)k}/(t-a_{J(1)})^{k(1)}$ for $i>1$. Thus

$$W_1(a_{J(1)})=(-1)^{k(1)\tau(1)}\rho W_2(a_{J(1)}),\quad \text{where } \rho>0$$

is some positive number. Similarly, $W_2(t)\neq 0$ and

$$W_2(a_{J(2)})=(-1)^{k(2)\tau(2)}\rho W_3(a_{J(2)}),\quad \text{where } \rho>0,$$

and W_3 has an obvious definition. This process continues, and we finally get

$$W_I(a_{J(I)})=(-1)^{k(I)\tau(I)}\rho, \quad \text{where } \rho>0.$$

Thus, we have

$$\operatorname{sgn} V=(-1)^{\tau}, \quad \text{where } \tau=\sum_{i=1}^{I} k(i)\tau(i).$$

This proves Theorem 2.2.

5. On Proposition 2.1. We shall use the boundary value problem adjoint to (1.1)–(1.2). In order to explain this, we associate a differential operator on $L^2(a, b)$ (called L again) with (1.1)–(1.2). Introduce the space $H_{n2}[a, b]=\{x(t): x\in C^{n-1}[a, b]$, $D^{n-1}x$ is absolutely continuous, and $D^n x\in L^2(a, b)\}$. The domain of the operator L is $D(L)=\{x(t)\in H_{n2}[a, b]: x$ satisfies the boundary conditions (1.2)$\}$. When (1.1) is disconjugate on $[a, b]$, the inverse L^{-1} of the operator L exists, is a bounded integral operator on $L^2(a, b)$, with the continuous kernel $G(t, s)$. The adjoint L^* of L is easy to describe if, for example, we assume that $p_k(t)\in C^k[a, b]$ in (2.11). In this case, the domain $D(L^*)$ of the adjoint L^* of (1.1)–(1.2) is the set of functions $x^*(t)$ which are piecewise smooth in the sense that $x^*\in H_{n2}(a_j, a_{j+1})$ for $j=1, \cdots, N-1$ and satisfies the boundary conditions

(5.1$_1$) $\quad D^{i-1}x^*(a_1)=0 \qquad$ for $1\leqq i\leqq n-m_1$,

(5.1$_k$) $\quad D^{i-1}x^*(a_k+0)=D^{i-1}x^*(a_k-0) \qquad$ for $1<k<N$, $1\leqq i\leqq n-m_k$,

(5.1$_N$) $\quad D^{i-1}x^*(a_N)=0 \qquad$ for $1\leqq i\leqq n-m_N$.

When (1.1), hence (2.11), is disconjugate on $[a, b]$ (Levin [7]; cf. Coppel [3, p. 104]), so that L and L^* have bounded inverses on $L^2(a, b)$, then $(L^*)^{-1}=(L^{-1})^*$ is an integral operator with kernel

$$G^*(t, s)=G(s, t). \tag{5.2}$$

That the domain $D(L^*)$ of L^* is as described above follows from the formal Green identity

$$\int_a^b (x^*Lx-xL^*x^*)\,dt=\sum_{j=1}^{N-1}\int_{a_j}^{a_{j+1}}$$
$$=\sum_{j=1}^{N-1}\left[\sum_{i=0}^{n-1}(-1)^iD^ix\sum_{k=i+1}^{n}(-1)^{k-1}D^{k-i-1}(p_kx^*)\right]_{a_j}^{a_{j+1}},$$

with $p_n\equiv 1$; that is, in terms of inner products on $L^2(a, b)$,

$$(Lx, x^*)-(x, L^*x^*)=\left[\sum_{i=0}^{n-1}(-1)^iD^ix\sum_{k=i+1}^{n}(-1)^{k-1}D^{k-i-1}(p_kx^*)\right]_a^b$$
$$-\sum_{j=2}^{N-1}\sum_{i=0}^{n-1}(-1)^iD^ix(a_j)\left[\sum_{k=i+1}^{n}(-1)^{k-1}D^{k-i-1}(p_kx^*)\right]_{a_j-0}^{a_j+0}.$$

Standard arguments show that the last equation is meaningful and that the right side is 0 for all $x\in D(L)$ if and only if $x^*\in D(L^*)$.

For fixed $s\in(a_j, a_{j+1})$ and $j=1, \cdots, N-1$, $x^*(t)=G^*(t, s)$, as a function of t, is a solution of $L^*x^*=0$ on (a_j, s), (s, a_{j+1}) and on (a_i, a_{i+1}) for $i\neq j$, satisfying the boundary conditions (5.1), and $D_t^{k-1}x^*(s+0)-D_t^{k-1}x^*(s-0)=0$ or 1 according as $1\leqq k<n$ or $k=n$.

By (2.6) and (5.2),

(5.3) $$y_j(t)=D_t^{m_j}G(t,s)|_{t=a_j,s=t}=D_s^{m_j}G^*(t,s)|_{s=a_j}.$$

In order to verify Proposition 2.1, put

$$Y^*(t,s)=G^*(t,s) \quad \text{for } a\leqq t<s, \qquad Z^*(t,s)=G^*(t,s) \quad \text{for } s<t\leqq b,$$

for fixed $s\in(a,b)$. Extend the definitions of Y^* and Z^* by the formula

$$Z^*(t,s)-Y^*(t,s)=X^*(t,s) \quad \text{for } a\leqq s,\ t\leqq b.$$

(Note that if $n=2$, then this extension of $x^*(t)=Y^*(t,s)$ on $s<t\leqq b$ is the solution of $L^*x^*=0$ satisfying the initial conditions $D^{i-1}x^*(s)=D^{i-1}Y^*(s-0,s)$ for $i=1,\cdots,n$. A similar remark applies to Z^*.)

For fixed $s\in(a_i,a_{i+1})$, $x^*(t)=Y^*(t,s)$ is a solution of $L^*x^*=0$ on each of the t-intervals (a_i,s), (s,a_{i+1}) and (a_k,a_{k+1}) for $k\neq i$, satisfying the boundary conditions

(5.4$_1$) $$D^{i-1}x^*(a_1)=0 \quad \text{for } 1\leqq i\leqq n-m_1,$$

(5.4$_k$) $$D^{i-1}x^*(a_k+0)=D^{i-1}x^*(a_k-0) \quad \text{for } 1<k<N,\ 1\leqq i\leqq n-m_k,$$

(5.4$_N$) $$D^{i-1}x^*(a_N)=-D_t^{i-1}X^*(b,s) \quad \text{for } 1\leqq i\leqq n-m_N.$$

This is clear from the boundary conditions satisfied by $Y^*=G^*$ for $a\leqq t<s$ and $Z^*=G^*=Y^*+X^*$ for $s<t\leqq b$.

It follows that $x^*(t)=D_s^{m_j}Y^*(t,s)$ is a solution of $L^*x^*=0$ on each of the intervals (a_i,s), (s,a_{i+1}) and (a_k,a_{k+1}) for $k\neq i$ satisfying

(5.5$_1$) $$D^{i-1}x^*(a_1)=0 \quad \text{for } 1\leqq i\leqq n-m_1,$$

(5.5$_k$) $$D^{i-1}x^*(a_k+0)=D^{i-1}x^*(a_k-0) \quad \text{for } 1<k<N,\ 1\leqq i\leqq n-m_k,$$

(5.5$_N$) $$D^{i-1}x^*(a_N)=-D_s^{m_j}D_t^{i-1}X^*(b,s) \quad \text{for } 1\leqq i\leqq n-m_N.$$

This characterizes the kernel $G_j^*(t,s)=D_s^{m_j}Y^*(t,s)$ for $a\leqq t<s$, $s\neq a_i$. Similarly, we can characterize $G_j^*=D_s^{m_j}Z^*(t,s)$ for $s<t\leqq b$, $s\neq a_i$. The last part of (5.3) and continuity considerations show that these characterizations for $a\leqq t<s$ and $s<t\leqq b$, when $s\neq a_i$, are also valid when $s=a_j$ for some a_j. Hence $s=a_j$ in (5.5) gives the existence of $y_j^0(t)$ in Proposition 2.1. Analogously, we obtain the existence of $y_j^1(t)$. This completes the proof.

REFERENCES

[1] P. W. BATES AND G. B. GUSTAFSON, *Green's function inequalities for two-point boundary value problems*, Pacific J. Math., 59 (1975), pp. 327–343.

[2] ———, *Maximization of Green's functions over classes of multipoint boundary value problems*, this Journal, 7 (1976), pp. 858–871.

[3] W. A. COPPEL, *Disconjugacy*, Lecture Notes in Mathematics, No. 220, Springer, New York, 1971.

[4] P. HARTMAN, *Ordinary Differential Equations*, S. M. Hartman, Baltimore, 1973.

[5] ———, *Unrestricted n-parameter families*, Rend. Circ. Mat. Palermo (2), 7 (1958), pp. 123–142.

[6] ———, *Principal solutions of disconjugate n-th order linear differential equations*, Amer. J. Math., 91 (1969), pp. 306–362 and 93 (1971), pp. 439–451.

[7] A. JU. LEVIN, *Some problems bearing on the oscillation of solutions of linear differential equations*, Dokl. Akad. Nauk SSSR, 148 (1963), pp. 512–515 = Soviet Math. Dokl., 4 (1963), pp. 121–124.

[8] Z. OPIAL, *On a theorem of Aramă*, J. Differential Equations, 3 (1967), pp. 88–91.

[9] JU. V. POKORNYĬ, *Some estimates of the Green's function of a multipoint boundary value problem*, Mat. Zametki, 4 (1968), pp. 533–540.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0005 $01.00/0

A STABILITY CRITERION OF DIAGONAL DOMINANCE TYPE*

F. NAKAJIMA†

Abstract. We consider a stability criterion for solutions of a linear system with coefficient matrix of diagonal dominance type. The result is applied to improve Fink's result [2] for the existence of almost periodic solutions of Liénard's equation.

1. Introduction. In a linear system

$$\dot{x} = A(t)x, \qquad x \in R^n, \quad \left(\cdot = \frac{d}{dt}\right) \tag{1.1}$$

let $A(t) = (a_{ij}(t))$, $i, j = 1, 2, \cdots, n$, be an $n \times n$ matrix of continuous functions for $-\infty < t < \infty$. It is known that the zero solution is uniformly asymptotically stable, if

$$a_{jj}(t) \leqq 0 \quad \text{for } 1 \leqq j \leqq n \quad \text{and} \quad -\infty < t < \infty \tag{1.2}$$

and if there is a constant $\delta > 0$ such that

$$\sum_{i=1\,(i \neq j)}^{n} |a_{ij}(t)| + \delta \leqq |a_{jj}(t)| \quad \text{for } 1 \leqq j \leqq n \quad \text{and} \quad -\infty < t < \infty \tag{1.3}$$

(cf. [1, p. 59]). As is stated in [5, p. 23], the condition (1.3) is called the strict diagonal dominance condition for $A(t)$ and it is a sufficient condition for

$$\det A(t) \neq 0. \tag{1.4}$$

A natural extension of (1.3) is the diagonal dominance condition that

$$\sum_{i=1\,(i \neq j)}^{n} |a_{ij}(t)| \leqq |a_{jj}(t)| \quad \text{for } 1 \leqq j \leqq n \quad \text{and} \quad -\infty < t < \infty. \tag{1.5}$$

In this paper we propose to extend the above stability criterion (1.2) and (1.3). In Theorem 1, we shall prove that the zero solution is uniformly asymptotically stable, if $A(t)$ is bounded on $(-\infty, \infty)$ and if conditions (1.2), (1.4) and (1.5) are satisfied. The restriction that $A(t)$ be bounded is essential for the theorem. But, system (1.1) with $A(t)$ bounded does arise naturally in many situations, for example, the case where $A(t)$ is periodic or almost periodic.

One great advantage of this extension will be seen in the following applications. In Theorem 2, we shall obtain a stability criterion for linear second-order scalar equation with variable coefficients. In Theorem 3, we shall prove the existence of almost periodic solutions for Liénard's equation with almost periodic forcing term.

Let R^n denote n-Euclidean space and set $R = (-\infty, \infty)$. For $x = (x_1, x_2, \cdots, x_n) \in R^n$, we define the norm of x by

$$|x| = \sum_{i=1}^{n} |x_i|.$$

Let $x(t, t_0, x_0)$ be the solution of system (1.1) through x_0 at t_0.

* Received by the editors July 21, 1976, and in revised form January 24, 1977.

† Mathematical Institute, Tôhoku University, Sendai, Japan. This work was supported in part by the Sakkokai Foundation.

DEFINITION 1. The zero solution is said to be *uniformly stable* if for each $\varepsilon > 0$, there exists a $\delta(\varepsilon) > 0$ such that

$$|x(t, t_0, x_0)| < \varepsilon \quad \text{for } t \geqq t_0$$

whenever $|x_0| < \delta(\varepsilon)$.

DEFINITION 2. The zero solution is said to be *uniformly asymptotically stable* if it is uniformly stable and if for each $\varepsilon > 0$, there exists a $T(\varepsilon) > 0$ such that

$$|x(t, t_0, x_0)| < \varepsilon \quad \text{for } t \geqq t_0 + T(\varepsilon)$$

whenever $|x_0| \leqq 1$.

Remark. In Definitions 1 and 2, if we are concerned only with $t_0 \geqq 0$, we say that the zero solution is uniformly stable for $t \geqq 0$ and is uniformly asymptotically stable for $t \geqq 0$, respectively.

2. Stability criterion.

THEOREM 1. *In system* (1.1), *let $A(t)$ be bounded on R. Assume that conditions* (1.2) *and* (1.5) *are satisfied for all $t \in R$ and that there is a constant $\alpha > 0$ such that*

$$|\det A(t)| \geqq \alpha \quad \textit{on } R.$$

Then the zero solution is uniformly asymptotically stable.

Remark. It is known that if $A(t)$ satisfies (1.3) with a constant $\delta > 0$, then $|\det A(t)| \geqq \delta^n$ (cf. [5, p. 16]).

To prove the theorem, we need the following lemmas.

DEFINITION 3. A square matrix A is said to be *irreducible* if A cannot be transformed to a matrix of the form

$$\begin{pmatrix} P & U \\ O & Q \end{pmatrix}$$

by permutation of indices, where P and Q are square matrices of order $\geqq 1$ and O is a zero matrix.

LEMMA 1. *If a square matrix A is irreducible and satisfies* (1.5) *and if for at least one j,*

$$\sum_{i=1\,(i \neq j)}^{n} |a_{ij}| < |a_{jj}|,$$

then A is nonsingular.

For the proof, see [5, p. 23].

LEMMA 2. *If a nonsingular $n \times n$ matrix $A = (a_{ij})$ satisfies* (1.5), *then all principal minors of A are nonsingular, namely,*

$$\det \begin{pmatrix} a_{j_1 j_1} & \cdots & a_{j_1 j_m} \\ a_{j_m j_1} & \cdots & a_{j_m j_m} \end{pmatrix} \neq 0 \quad \text{for } 1 \leqq j_1 < j_2 < \cdots < j_m \leqq n.$$

Proof of Lemma 2. Let A_1 be an $m \times m$ $(m < n)$ principal minor of A. Then, for a permutation matrix Q,

$$QAQ^T = \begin{pmatrix} A_1 & * \\ A_2 & * \end{pmatrix},$$

where A_2 has $(n-m)$ rows and m columns and Q^T denotes the transposed matrix of Q. Moreover, from the definition of irreducibility, we can choose an $m \times m$ permutation matrix Q_m such that

$$Q_m A_1 Q_m^T = \begin{pmatrix} B_1 & C_2 & & \\ & B_2 & \ddots & \\ & & \ddots & C_p \\ & O & & B_p \end{pmatrix}, \tag{2.1}$$

where B_i is an $r_i \times r_i$ irreducible matrix, $\sum_{i=1}^{p} r_i = m$, and C_i has $(r_1 + r_2 + \cdots + r_{i-1})$ rows and r_i columns for $2 \leqq i \leqq p$. In particular, in the case where A_1 is irreducible, B_1 must be A_1 itself, and the matrices $B_2, B_3, \cdots, B_p, C_2, C_3, \cdots, C_p$ are not present.

Setting $P = (Q_m \oplus I)Q$ for $(n-m) \times (n-m)$ unit matrix I, we have

$$B = PAP^T = \begin{pmatrix} Q_m A_1 Q_m^T & * \\ A_2 Q_m^T & * \end{pmatrix} = \left(\begin{array}{cccc|c} B_1 & C_2 & & & \\ & B_2 & \ddots & & * \\ & & \ddots & C_p & \\ & O & & B_p & \\ \hline D_1 & D_2 & \cdots & D_p & * \end{array}\right),$$

where $A_2 Q_m^T = (D_1, D_2, \cdots, D_p)$ and D_i has $(n-m)$ rows and r_i columns. Since the diagonal dominance condition (1.5) is invariant under the permutation of indices, B also satisfies (1.5). Hence, letting

$$B_i = (b_{jk}), \quad C_i = (c_{jk}), \quad D_i = (d_{jk})$$

for a fixed i, $1 \leqq i \leqq p$, we have

$$\sum_{j=1\,(j \neq k)}^{r_i} |b_{jk}| + \sum_j |c_{jk}| + \sum_j |d_{jk}| \leqq |b_{kk}| \qquad (1 \leqq k \leqq r_i), \tag{2.2}$$

where the summations on j are taken along columns and $C_1 = O$ for convenience.

If $C_i \neq O$ or $D_i \neq O$, then

$$\sum_j |c_{jk}| + \sum_j |d_{jk}| > 0 \quad \text{for some } k, \quad 1 \leqq k \leqq r_i,$$

and hence for this k,

$$\sum_{j=1\,(j \neq k)}^{r_i} |b_{jk}| < |b_{kk}|$$

by (2.2). Therefore it follows from Lemma 1 that

$$\det B_i \neq 0, \tag{2.3}$$

since B_i is irreducible. If $C_i = O$ and $D_i = O$, then we have the form of

$$\det B = \det B_i \times (\cdots)$$

which also implies (2.3), because $\det B \neq 0$. In any case, we have $\det B_i \neq 0$. Since these are true for all i, $1 \leqq i \leqq p$, it follows from (2.1) that

$$\det A_1 \neq 0.$$

LEMMA 3. *If system* (1.1) *satisfies conditions* (1.2) *and* (1.5), *then the norm of solution* $x(t)$, $|x(t)|=\sum_{i=1}^{n}|x_i(t)|$, *is nonincreasing, and consequently the zero solution is uniformly stable.*

For the proof, see [1, p. 59].

Proof of Theorem 1. As is stated in Lemma 3, the zero solution is uniformly stable, and hence it is sufficient to show that for each $\varepsilon>0$ there exists a $T(\varepsilon)>0$ such that

$$|x(t, t_0, x_0)|<\varepsilon \quad \text{at some } t, \quad t_0 \leqq t \leqq t_0+T(\varepsilon),$$

whenever $|x_0|\leqq 1$.

Suppose that this is not true. Then there exists a constant $\varepsilon>0$, a sequence of solutions $\{x(t, k)\}$ and a sequence $\{t_k\}$ such that

$$|x(t, k)|\geqq \varepsilon \quad \text{on } t_k \leqq t \leqq t_k+k^2$$

and

$$|x(t_k, k)|\leqq 1.$$

Since $|x(t, k)|$ is nonincreasing, we have

$$\varepsilon \leqq |x(t, k)| \leqq 1 \quad \text{on } t_k \leqq t \leqq t_k+k^2$$

and there exists a subinterval $[s_k, s_k+k]$ of $[t_k, t_k+k^2]$ such that

$$\Big||x(t, k)|-|x(t', k)|\Big|<\frac{1}{k} \quad \text{for } s_k \leqq t, \quad t' \leqq s_k+k.$$

Setting $\varphi(t, k)=x(t+s_k, k)$, we obtain

$$\dot{\varphi}(t, k)=A(t+s_k)\varphi(t, k), \tag{2.4}$$

$$\varepsilon \leqq |\varphi(t, k)| \leqq 1 \quad \text{on } 0 \leqq t \leqq k, \tag{2.5}$$

$$\Big||\varphi(t, k)|-|\varphi(t', k)|\Big|<\frac{1}{k} \quad \text{for } 0 \leqq t, \quad t' \leqq k. \tag{2.6}$$

Since $A(t)$ is bounded, it follows from (2.4) and (2.5) that $\{\varphi(t, k)\}$ is uniformly bounded and equi-continuous on each finite interval of R, and thus, by Ascoli–Arzela's theorem, $\{\varphi(t, k)\}$ can be assumed to converge uniformly on each finite interval of R.

Defining $y(t)$ by

$$y(t)=\lim_{k\to\infty}\varphi(t, k),$$

it follows from (2.5) and (2.6) that there is a constant $\beta>0$, $\varepsilon \leqq \beta \leqq 1$, such that

$$|y(t)|=\beta \quad \text{for } t \geqq 0.$$

Since $y(t)=(y_1(t), y_2(t), \cdots, y_n(t))$ is continuous on R, we can choose an interval $I=[\sigma_1, \sigma_2]$ $(0<\sigma_1<\sigma_2)$ such that

$$y_{j_1}(t), y_{j_2}(t), \cdots, y_{j_h}(t)>0 \quad \text{on } I,$$

$$y_{j_{h+1}}(t), y_{j_{h+2}}(t), \cdots, y_{j_m}(t)<0 \quad \text{on } I$$

and

$$y_{j_{m+1}}(t) \equiv y_{j_{m+2}}(t) \equiv \cdots \equiv y_{j_n}(t) \equiv 0 \quad \text{on } I.$$

Here we note that $\{j_1, j_2, \cdots, j_m\} \neq \varnothing$ because $y(t) \not\equiv 0$. Then

$$|y(t)| = \sum_{p=1}^{h} y_{j_p}(t) - \sum_{p=h+1}^{m} y_{j_p}(t) = \beta \quad \text{on } I.$$

Letting

$$f(t, k) = \sum_{p=1}^{h} \varphi_{j_p}(t, k) - \sum_{p=h+1}^{m} \varphi_{j_p}(t, k) \tag{2.7}$$

we have

$$\lim_{k\to\infty} f(t, k) = |y(t)| = \beta \quad \text{on } I,$$

and there is a sequence $\{\theta_k\} \subset I$ such that

$$\lim_{k\to\infty} \dot{f}(\theta_k, k) = 0. \tag{2.8}$$

Moreover, since I is compact and $A(t)$ is bounded on R, we can assume that

$$\lim_{k\to\infty} \theta_k = \theta \quad \text{for some } \theta \in I$$

and

$$\lim_{k\to\infty} A(s_k + \theta_k) = B \quad \text{for some } n \times n \text{ matrix } B.$$

Clearly B satisfies (1.2), (1.5) and $|\det B| \geqq \alpha$.

Differentiating both sides of (2.7) at $t = \theta_k$ and using relation (2.4), we find

$$\dot{f}(\theta_k, k) = \sum_{p=1}^{h} \left(\sum_{q=1}^{n} a_{j_p j_q}(s_k + \theta_k)\varphi_{j_q}(\theta_k, k) \right) - \sum_{p=h+1}^{m} \left(\sum_{q=1}^{n} a_{j_p j_q}(s_k + \theta_k)\varphi_{j_q}(\theta_k, k) \right),$$

and

$$\lim_{k\to\infty} \dot{f}(\theta_k, k) = \sum_{p=1}^{h} \left(\sum_{q=1}^{m} b_{j_p j_q} y_{j_q}(\theta) \right) - \sum_{p=h+1}^{m} \left(\sum_{q=1}^{m} b_{j_p j_q} y_{j_q}(\theta) \right),$$

where $B = (b_{ij})$, since we have

$$\lim_{k\to\infty} \varphi_j(\theta_k, k) = y_j(\theta) \quad \text{for } 1 \leqq j \leqq n$$

and

$$y_{j_q}(\theta) = 0 \quad \text{for } m + 1 \leqq q \leqq n.$$

By (2.8), we have

$$0=\sum_{q=1}^{m}\left(\sum_{p=1}^{h}b_{j_pj_q}-\sum_{p=h+1}^{m}b_{j_pj_q}\right)y_{j_q}(\theta). \tag{2.9}$$

Since B satisfies (1.2) and (1.5),

$$\sum_{p=1}^{h}b_{j_pj_q}-\sum_{p=h+1}^{m}b_{j_pj_q}\begin{cases}\leqq 0 & \text{for } 1\leqq q\leqq h,\\ \geqq 0 & \text{for } h+1\leqq q\leqq m\end{cases}$$

Therefore each term of right hand side of (2.9) is nonpositive, because we have

$$y_{j_q}(\theta)\begin{cases}>0 & \text{for } 1\leqq q\leqq h,\\ <0 & \text{for } h+1\leqq q\leqq m.\end{cases}$$

Then it follows from (2.9) that

$$\left(\sum_{p=1}^{h}b_{j_pj_q}-\sum_{p=h+1}^{m}b_{j_pj_q}\right)y_{j_q}(\theta)=0 \quad \text{for } 1\leqq q\leqq m,$$

which implies

$$\sum_{p=1}^{h}b_{j_pj_q}-\sum_{p=h+1}^{m}b_{j_pj_q}=0 \quad \text{for } 1\leqq q\leqq m$$

since $|y_{j_q}(\theta)|\neq 0$ for $1\leqq q\leqq m$. Thus we have

$$\det\begin{pmatrix}b_{j_1j_1} & \cdots & b_{j_1j_m}\\ b_{j_mj_1} & \cdots & b_{j_mj_m}\end{pmatrix}=0. \tag{2.10}$$

On the other hand, B satisfies (1.5) and $|\det B|\geqq\alpha>0$, and thus it follows from Lemma 2 that all principal minors of B are nonsingular, which contradicts (2.10). This proves that the zero solution of system (1.1) is uniformly asymptotically stable.

COROLLARY. *If system* (1.1) *is defined only for* $t\geqq 0$ *and all assumptions of Theorem* 1 *are satisfied for* $t\geqq 0$, *then the zero solution is uniformly asymptotically stable for* $t\geqq 0$.

Proof. We construct the system defined on R by

$$\dot{x}=A_0(t)x, \qquad x\in R^n, \tag{2.11}$$

where

$$A_0(t)=\begin{cases}A(t) & \text{for } t\geqq 0\\ A(0) & \text{for } t<0.\end{cases}$$

Since system (2.11) satisfies all assumptions of Theorem 1 on R, the zero solution is uniformly asymptotically stable on R, and furthermore, since system (1.1) coincides with system (2.11) for $t\geqq 0$, this proves our conclusion.

3. Application 1. We now apply Theorem 1 to the second-order scalar equation

$$\ddot{x}+a(t)\dot{x}+b(t)x=0, \qquad x\in R, \tag{3.1}$$

where $a(t)$ and $b(t)$ are continuous on R.

THEOREM 2. *Let $a(t)$ and $b(t)$ be positive and bounded on R. Assume that there exist positive constants α_1, α_2, β_1, β_2 such that*

(i) $\alpha_1 \leqq a(t) \leqq \alpha_2$, $\beta_1 \leqq b(t) \leqq \beta_2$ *on* R,

(ii) $3\alpha_1 > \alpha_2$,

and

(iii) $(3\alpha_1 - \alpha_2)(\alpha_1 + \alpha_2) \geqq 8\beta_2$.

Then the zero solution of (3.1) *is uniformly asymptotically stable, and consequently, for any solution* $x(t)$,

$$|x(t)| + |\dot{x}(t)| \to 0 \quad \textit{as } t \to \infty.$$

Proof. First of all we shall show that there is a constant $\lambda > 0$ such that

$$\lambda^2 - 2a(t)\lambda + 2b(t) \leqq 0 \quad \text{on } R, \tag{3.2}$$

which is equivalent to showing that

$$\inf_{t\in R} \left(a(t) + \sqrt{a^2(t) - 2b(t)}\right) \geqq \sup_{t \in R} \left(a(t) - \sqrt{a^2(t) - 2b(t)}\right) > 0. \tag{3.3}$$

From conditions (ii) and (iii), we have

$$2\sqrt{\alpha_1^2 - 2\beta_2} \geqq \alpha_2 - \alpha_1$$

and hence we have

$$2 \inf_{t \in R} \sqrt{a^2(t) - 2b(t)} \geqq \sup_{t\in R} a(t) - \inf_{t \in R} a(t)$$

which implies (3.3).

For this λ, let

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \lambda & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ \dot{x} \end{pmatrix}.$$

Then equation (3.1) is equivalent to the system

$$\begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = A(t) \begin{pmatrix} u \\ v \end{pmatrix}, \tag{3.4}$$

where

$$A(t) = \frac{1}{\lambda} \begin{pmatrix} -b(t) & \lambda^2 - a(t)\lambda + b(t) \\ -b(t) & -a(t)\lambda + b(t) \end{pmatrix}.$$

We shall verify that $A(t)$ satisfies all assumptions in Theorem 1. It is clear that $A(t)$ is bounded on R and $\det A(t) = b(t) \geqq \beta_1$. The conditions (1.2) and (1.5) require that

$$-b(t) + |-b(t)| \leqq 0 \tag{3.5}$$

and

$$-a(t)\lambda + b(t) + |\lambda^2 - a(t)\lambda + b(t)| \leqq 0. \tag{3.6}$$

Condition (3.5) is trivial and (3.6) is equivalent to (3.2). Hence, by Theorem 1, the zero solution of (3.4) is uniformly asymptotically stable, and consequently the zero solution of (3.1) also is uniformly asymptotically stable and

$$|x(t)| + |\dot{x}(t)| \to 0 \quad \text{as } t \to \infty.$$

4. Application 2. We shall consider the existence of almost periodic solutions of Liénard's equation

$$\ddot{x}+f(x)\dot{x}+g(x)=e(t), \qquad x\in R, \tag{4.1}$$

where $e(t)$ is almost periodic in t.

THEOREM 3. *Let $f(x)$ be continuous and $g(x)$ be once differentiable. Assume that there exist positive constants α, β_1, β_2 such that*

$$f(x)\geqq\alpha, \qquad \beta_2\geqq\frac{dg}{dx}(x)\geqq\beta_1 \quad \text{for } x\in R$$

and

$$\alpha^2\geqq 2\beta_2.$$

Then there exists an almost periodic solution $p(t)$ which is unique and whose module is contained in the module of $e(t)$. Moreover, for any solution $x(t)$,

$$|x(t)-p(t)|+|\dot{x}(t)-\dot{p}(t)|\to 0 \quad \text{as } t\to\infty.$$

Remark. The case where $\alpha^2>2\beta_2$ was discussed by Fink [2]. Theorem 3 contains the case where $\alpha^2=2\beta_2$, and furthermore a perturbation method enables us to permit $\alpha^2<2\beta_2$, if $|\alpha^2-2\beta_2|$ is sufficiently small.

First of all we shall state the existence theorem of bounded solutions of equation (4.1) due to Loud [3].

LEMMA 4. *In equation* (4.1), *suppose that*

$$f(x)\geqq\alpha>0 \quad \text{and} \quad \frac{dg}{dx}(x)\geqq\beta_1>0 \quad \text{on } R.$$

Then all solutions $x(t)$ ultimately satisfy

$$|x(t)|\leqq\min\left\{\frac{E}{\beta_1}+\frac{4E}{\alpha^2}, \frac{E}{\beta_1}+\frac{4E}{\beta_1^{1/2}\alpha}\right\},$$

$$|\dot{x}(t)|\leqq\frac{4E}{\alpha},$$

where $E=\sup_{t\in R}|e(t)|$.

It is easily seen that the following lemma is a special case of Theorem 3 in [4].

LEMMA 5. *In the system*

$$\dot{x}=F(t,x), \qquad x\in R^n,$$

let $F(t,x)$ be almost periodic in t uniformly for $x\in R^n$ and for each $r>0$ let there exist a constant $L=L(r)>0$ such that

$$|F(t,x)-F(t,y)|\leqq L|x-y| \quad \text{for } |x|, |y|\leqq r \quad \text{and} \quad t\in R.$$

If $x(t)$ is a bounded solution on R and if for any solution $y(t)$, $|x(t)-y(t)|$ is monotone decreasing to zero as $t\to\infty$, then $x(t)$ is a unique almost periodic solution and its module is contained in the module of $F(t,x)$.

Proof of Theorem 3. The proof is similar to Theorem 2 in [2] which used conditions (1.2) and (1.3) as a stability criterion, but we shall employ Corollary 1 instead of them. We first let

$$u=x, \qquad v=(\dot{x}+F(x))\times\frac{2}{\alpha}, \tag{4.2}$$

where

$$F(x)=\int_0^x f(s)\,ds-\frac{\alpha}{2}x.$$

Then equation (4.1) is equivalent to the system

$$\text{(4.3)}\qquad \begin{aligned}\dot u&=-F(u)+\frac{\alpha}{2}v,\\ \dot v&=F(u)-\frac{2}{\alpha}g(u)-\frac{\alpha}{2}v+\frac{2}{\alpha}e(t).\end{aligned}$$

All solutions of (4.3) are bounded in the future, because all solutions of (4.1) are bounded in the future together with their derivatives by Lemma 4. Hence it follows from the almost periodicity of $e(t)$ that there is at least one bounded solution on R, say $(u_0(t), v_0(t))$.

Considering the difference between $(u_0(t), v_0(t))$ and any solution $(u(t), v(t))$ of (4.3)

$$w(t)=u_0(t)-u(t),\ z(t)=v_0(t)-v(t),$$

we get the variational equation

$$\text{(4.4)}\qquad \begin{pmatrix}\dot w\\ \dot z\end{pmatrix}=A(t)\begin{pmatrix}w\\ z\end{pmatrix},$$

where

$$A(t)=\begin{pmatrix}-p(t) & \dfrac{\alpha}{2}\\ p(t)-\dfrac{2}{\alpha}q(t) & -\dfrac{\alpha}{2}\end{pmatrix},$$

$$p(t)=\int_0^1 f((1-s)u(t)+su_0(t))\,ds-\frac{\alpha}{2}$$

and

$$q(t)=\int_0^1 \frac{dg}{dx}((1-s)u(t)+su_0(t))\,ds.$$

Clearly $p(t)\geqq\alpha/2$ and $\beta_1\leqq q(t)\leqq\beta_2$.

We shall verify that $A(t)$ satisfies all assumptions in Corollary 1. First of all, $A(t)$ is bounded in the future, because $(u_0(t), v_0(t))$ and $(u(t), v(t))$ are bounded in the future. It is clear that the diagonal elements of $A(t)$ are negative and

$$\det A(t)=q(t)\geqq\beta_1\quad\text{on } R.$$

The diagonal dominance condition (1.5) for $A(t)$ requires that

$$-p(t)+\left|p(t)-\frac{2}{\alpha}q(t)\right|\leqq 0,\qquad -\frac{\alpha}{2}+\left|-\frac{\alpha}{2}\right|\leqq 0,$$

which is equivalent to

$$q(t)\leqq\alpha p(t)$$

and this is satisfied by

$$q(t) \leqq \beta_2 \leqq \alpha^2/2 \leqq \alpha p(t).$$

Therefore, by Theorem 1 the zero solution of (4.4) is uniformly asymptotically stable and

$$|u_0(t) - u(t)| + |v_0(t) - v(t)| \to 0 \quad \text{as } t \to \infty,$$

where the convergence is monotone decreasing by Lemma 3. Thus, applying Lemma 5 to system (4.3), we find that there exists a unique almost periodic solution with the module contained in the module of $e(t)$. From the property of transformation (4.2), this proves our conclusion for equation (4.1).

REFERENCES

[1] W. A. COPPEL, *Stability and Asymptotic Behavior of Differential Equations*, Heath Math. Monograph, 1965.

[2] A. M. FINK, *Convergence and almost periodicity of solutions of forced Liénard equations*, SIAM J. Appl. Math., 26 (1974), pp. 26–34.

[3] W. S. LOUD, *Boundedness and convergence of solutions of* $\ddot{x} + c\dot{x} + g(x) = e(t)$, Duke Math. J., 24 (1957), pp. 63–72.

[4] G. SEIFERT, *Almost periodic solutions and asymptotic stability*, J. Math. Anal. Appl., 21 (1968), pp. 136–149.

[5] R. S. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0006 $01.00/0

A NOTE ON THE RAYLEIGH POLYNOMIALS*

E. C. OBI†

Abstract. Some formulae and representations of the nth order Rayleigh functions, σ_n, and their associated polynomials, ϕ_n, are used to determine polynomials which dominate, or are dominated by, ϕ_n, and to discuss improved lower bounds for their real roots. In the process the ties between ϕ_n and the Catalan numbers are enhanced.

1. Introduction. The nth order Rayleigh polynomials, ϕ_n, arises as explained below. Consider the infinitely many zeros, $j_{\nu,m}$, $m = 1, 2, 3, \cdots$, of $z^{-\nu} J_\nu(z)$, ordered by $|\mathrm{Re}\,(j_{\nu,m})| < |\mathrm{Re}\,(j_{\nu,m+1})|$, where J_ν is the Bessel function of order ν. On letting

$$\sigma_n(\nu) = \sum_{m=1}^{\infty} (j_{\nu,m})^{-2n}, \qquad n = 1, 2, 3, \cdots, \tag{1}$$

and expressing J_ν in its Weierstrass infinite product expansion using the zeros, $j_{\nu,m}$, we find the following series derived in [1, pp. 528–529]:

$$\tfrac{1}{2} z J_{\nu+1}(z) = J_\nu(z) \sum_{n=1}^{\infty} \sigma_n(\nu) z^{2n}.$$

If we replace $J_{\nu+1}$ and J_ν with their own power series and equate corresponding coefficients of z^{2n} on both sides of the above equation, we find

$$\sum_{k=1}^{n} (-1)^{k-1} 4^{\kappa} (k!)^2 \binom{n}{k} \binom{\nu+n}{k} \sigma_n(\nu) = n. \tag{*}$$

Also, by using the well-known derivative formulae connecting J_ν with $J_{\nu+1}$, it can be shown [1, (22)] that

$$(\nu+n)\sigma_n(\nu) = \sum_{k=1}^{n-1} \sigma_k(\nu)\sigma_{n-k}(\nu) \qquad (n > 1). \tag{**}$$

Now putting $n = 1$ in (*) we obtain $\sigma_1(\nu) = \frac{1}{4}(\nu+1)$. Using this and letting $n = 2$ in (**) we get the formula for $\sigma_2(\nu)$. In this way formulae for successive orders, n, may be obtained inductively. (The first few of these are listed at the end of article for reference purposes.) Thus, for each n, $\sigma_n(\nu)$ is a rational function of ν, viz. $\sigma_n(\nu) = \phi_n(\nu)/(\pi_n(\nu))$ where ϕ_n, π_n are polynomials.

This rational function is called the Rayleigh function of order n. Its denominator, Π_n, may always be calculated by the simple algebraic formula, $\Pi_n(\nu) = 4^n \prod_{k=1}^{n} (\nu+k)^{[n/k]}$, where $[x]$ denotes the greatest integer function. But a simple algebraic formula for the numerator, ϕ_n (and hence for σ_n itself) is unknown. This $\phi_n(\nu)$ is called the nth order Rayleigh polynomial in ν. Among its many amazing properties are that its leading coefficient, e_n, for each n, is always the nth Catalan number (see § 2 and last paragraph of this paper) and if $d(n)$ is the degree of $\phi_n(\nu)$, then $d(n) - d(n-1)$ provides a formula for number of the nontrivial divisors of n [2, (7)]. Recent studies of $\sigma_n(\nu)$ and $\phi_n(\nu)$ yielding representation formulae and ties with Riccati differential equations, etc., were carried out by Nand Kishore in [1]–[5]. Many analytic properties of the Rayleigh function and $\sigma_n^{(m)}(\nu)$ are given in [7]. Rayleigh, Lamb, Airey and others (cf. [8, p. 502]) had used the function to study and

* Received by the editors April 30, 1976, and in revised form September 16, 1976.
† Department of Mathematics, University of Nigeria, Nsukka Campus, East Central State, Nigeria.

deduce the smallest positive zeros of J_ν, for various orders, ν. In view of the definitions above, we will sometimes find it convenient to use ϕ_n in the form

$$\phi_n(\nu)=4^n\prod_{k=1}^{n}(\nu+k)^{[n/k]}\sigma_n(\nu). \tag{1a}$$

In what follows we discuss a number of facts about the properties of ϕ_n and an improved lower bound for its real roots.

2. The polynomial on $(-\infty, -n+1)$. In [2, § 5] it is shown by Kishore that $-n$ is a lower bound for the real zeros of ϕ_n, and that all these zeros lie within $(-n, -2)$. In this section we will narrow this interval first to $(-n+1, -2)$ and then discuss the possibility of further improvement on the lower bound by as many as $n-[n/2]-1$ units.

Let

$$e_n=\frac{1}{n}\binom{2n-2}{n-1}, \qquad d(n)=\deg\phi_n(\nu),$$

and, by a theorem of Kishore [3, p. 515], write (1a) in the form

$$\phi_n(\nu)=\sum_{i=1}^{c(n)}2^{n_i}\prod_{j=2}^{n-1}(\nu+j)^{n_{ij}} \tag{2}$$

where (2i) at most one $n_i=0$;

$$\sum_{i=1}^{c(n)}2^{n_i}=e_n; \qquad \sum_{j=2}^{n-1}n_{ij}=d(n)=1-2n+\sum_{s=1}^{n}[n/s] \tag{2ii}$$

for each $i=1, 2, \cdots, c(n)$;

and (2iii) for each integer $s\in(1, n)$, $n>3$, there is a summand in (2) with $-s$ as a zero of multiplicity $\leqq[n/s]$ (i.e. there is i such that $0<n_{is}\leqq[n/s]$). As before $[x]$ stands for the greatest integer function, and the number $c(n)$ satisfies

$$c(n)=\sum_{k=1}^{[n/2]}c(k)c(n-k), \quad \text{with } c(1)=1.$$

The number e_n, is called the nth Catalan number.[1]

Now let ε be fixed in $0\leqq\varepsilon<1$, and consider the interval $\nu\leqq-n+\varepsilon$. If $d(n)$ is even, then on rewriting (2) in the form

$$\phi_n(\nu)=\sum_{i=1}^{c}2^{n_i}\prod_{s=1}^{d(n)}(\nu+b_{is}), \quad 2\leqq b_{is}\leqq n-1, \quad c=c(n), \tag{2a}$$

we have on $\nu\leqq-n+\varepsilon$, that

$$\phi_n(\nu)=\sum_{i=1}^{c}2^{n_i}\prod_{s=1}^{d(n)}-(\nu+b_{is})\geqq\sum_{i=1}^{c}2^{n_i}\prod_{s=1}^{d(n)}(1-\varepsilon)=e_n(1-\varepsilon)^{d(n)}>0.$$

On the other hand, if $d(n)$ is odd, then on $\nu\leqq-n+\varepsilon$,

$$\begin{aligned}\phi_n(\nu)=-\sum_{i=1}^{c}2^{n_i}\prod_{s=1}^{d(n)}-(\nu+b_{is})&\leqq-\sum_{i=1}^{c}2^{n_i}\prod_{s=1}^{d(n)}(1-\varepsilon)\\&=-e_n(1-\varepsilon)^{d(n)}<0.\end{aligned}$$

[1] Well-known in combinatorial mathematics, e_n has many other noteworthy properties which we mention at end of article. Meanwhile, due to (2ii), the leading coefficient of each ϕ_n is equal to e_n. See also (4a) and (6f).

Thus ϕ_n never vanishes on $(-\infty, -n+1)$. Hence all real roots of ϕ_n lie somewhere within $[-n+1, -2)$. Let $w_i(\nu) \equiv w_{n,i}(\nu)$ be the product $\prod$ in (2a). Considering $\log w_i$, we have

(3)
$$\phi_n'(\nu) = \sum_{i=1}^{c} \sum_{s=1}^{d(n)} (2^{n_i} w_i(\nu))/(\nu + b_{is}),$$
$$\phi_n''(\nu) = \sum_{i=1}^{c} \sum_{s=1}^{d(n)} 2^{n_i} \frac{w_i\left[(\nu+b_{is})\sum_{r=1}^{d(n)} \dfrac{1}{\nu+b_{ir}} - 1\right]}{(\nu+b_{is})^2}.$$

But if $d(n)$ is even, $w_i(\nu) > 0$ on $(-\infty, -n+1)$, and so

(3a) $$\phi_n'(\nu) < 0 \quad \text{on } (-\infty, -n+1);$$

while if $d(n)$ is odd,

(3b) $$\phi_n'(\nu) > 0 \quad \text{on } (-\infty, -n+1).$$

If $n = 4$ or 5, $d(n) = 1$ and ϕ_n is linear in ν, and hence ϕ_n'' vanishes. The following lemma helps to test the sign of ϕ_n'' when $n > 5$.

LEMMA 1. $d(n) > 1$ *iff* $n > 5$.

Proof. Since $d(n) > 1$ iff $[n/2] + \cdots + [n/n] > n$, the sufficiency follows easily by induction. The necessity holds since if $n \leqq 5$, $d(n) = 0$ or 1, in contradiction.

Now since $\nu + b_{ir} < 0$ on $(-\infty, -n+1)$ for each $r = 1, 2, \cdots, d(n)$, the lemma implies that

$$(\nu + b_{is}) \sum_{r=1}^{d(n)} \frac{1}{\nu + b_{ir}} > 1 \quad \text{when } n > 5.$$

Thus, since w_i is positive or negative on $(-\infty, -n+1)$ according as $d(n)$ is even or odd, it follows that

(3c) $$\phi_n''(\nu) \text{ is positive or negative on } (-\infty, -n+1)$$
$$\text{according as } d(n) \text{ is even or odd, if } n > 5.$$

We exhibit an inequality between $\phi_n(\nu)$ and $e_n\{-(\nu+2)\}^{d(n)}$ on $(-\infty, -n+1]$. If we let $\varepsilon \to 1$ following (2a), $\phi_n(\nu) \geqq 0$ on $(-\infty, -n+1]$ and we also have, if $d(n)$ is even,

$$\phi_n(\nu) = \sum_{i=1}^{c} \left\{ 2^{n_i} \prod_{s=1}^{d(n)} -(\nu + b_{is}) \right\} \leqq \sum_{i=1}^{c} \left\{ 2^{n_i} \prod_{s=1}^{d(n)} -(\nu+2) \right\}$$
$$= e_n\{-(\nu+2)\}^{d(n)}.$$

If $d(n)$ is odd, then similarly $\phi_n(\nu) \leqq 0$ on $\nu \leqq -n+1$ and we also have $\phi_n(\nu) \geqq -e_n\{-(\nu+2)\}^{d(n)}$.

Noting that $d(n)$ is even iff $\sum_{s=1}^{n-1} [n/s]$ is even, we summarize as follows:

(4a) *If* $\sum_{k=1}^{n-1} [n/k]$ *is even, then on* $(-\infty, -n+\varepsilon)$, $0 \leqq \varepsilon < 1$, *the Rayleigh polynomial,* $\phi_n(\nu)$, *is a positive strictly decreasing function which lies above the line* $y = e_n(1-\varepsilon)^{d(n)}$, *and is dominated by the polynomial* $e_n\{-(\nu+2)\}^{d(n)}$. *Furthermore, in view of* (3c), *the concavity of* ϕ_n *is upward* $(n > 5)$ *in the interval*;

(4b) *If* $\sum_{k=1}^{n-1} [n/k]$ *is odd, then on* $(-\infty, -n+\varepsilon)$, $0 \leqq \varepsilon < 1$, *the Rayleigh polynomial,* $\phi_n(\nu)$, *is a negative strictly increasing function which lies below the line* $y =$

$-e_n(1-\varepsilon)^{d(n)}$, *and dominates the polynomial* $-e_n\{-(\nu+2)\}^{d(n)}$. *Furthermore, in view of* (3c), *the concavity of* $\phi_n (n>5)$ *is downward in the interval*;

(4c) *If* ν_0 *is a real zero of* ϕ_n, *then* $-n+1 \leqq \nu_0 < -2$.

The following lemma leads to a further narrowing of the interval of (4c) for the real nonintegral roots.

LEMMA 2 (Kishore [6]). (a) *The Rayleigh function*, $\sigma_n(\nu)$, *given in* (1), *is positive on* $(-n, -n+1)$ *for all* n;

(b) *If* $\nu \leqq -n-2$, *and* $2 \leqq m \leqq n$, *then* $(\sigma_m(\nu)/\sigma_{n-1}(\nu)) < \{m/(2n(n+4))\}$;

(c) $\sigma_n(\nu) < 0$ *on* $(-n+1, -n+2)$ *if* $n>4$; *and*

(d) $\sigma_n(\nu) > 0$ *on* $(-n+2, -n+3)$ *if* $n>5$.

We give a proof of (a) by a short method. (The proofs of (b), (c) and (d), due to Professor Kishore, are quite lengthy, but an outline is given for convenience.) To prove (a), we will first note the following main aspects of a representation for $\sigma_n(\nu)$. Let us express the function in the form

(5) $$\sigma_n(\nu) = \sum_{i=1}^{c(n)} 2^{-N_i} P_i^{-1}(\nu) \qquad \text{(see [3, (20)])}$$

where $N_i = 2n - n_i > 0$. $P_i(\nu)$, obtained in the proof of (5) (see [3, (19), (20)]), is the polynomial

(5a) $$P_i(\nu) = \prod_{j=1}^{n} (\nu+j)^{[n/j]-n_{ij}}$$

where n_{ij} is as in (2), with

(5b) $$n_{i1} = n_{in} = 0 \quad \text{and} \quad [n/j] - n_{ij} \geqq 0,$$

due to (2) and (2iii). The degree of P_i is now seen easily, in view of (2ii), to be $2n-1$. It is also clear now that P_i has positive integral coefficients and only negative integral zeros lying in $[-n, -1]$, and that one may write P_i in either of the forms

(5c) $$P_i(\nu) = \prod_{s=1}^{m} (\nu+b_{is})^{a_{is}}$$

where $1 \leqq b_{is} \leqq n$, $1 \leqq a_{is} \leqq n$, $\sum_{s=1}^{m} a_{is} = 2n-1$, $1 \leqq m \leqq n$; or

(5d) $$P_i(\nu) = \prod_{s=1}^{2n-1} (\nu+b_{is}), \quad 1 \leqq b_{is} \leqq n.$$

We remark that the most important thing about $\{P_i\}_{i=1}^{c(n)}$ for our purposes is that $\sum_{s=1}^{m} a_{is} = 2n-1$, i.e. the P_i's have a common degree. Also, putting (5b) into (5a), we see that P_i always retains the factors $(\nu+1)$ n fold, and $(\nu+n)$ one fold. As usual an n fold factor is counted as n factors.

Proof of Lemma 2(a). Take (5a, d). Since $(\nu+n)$ is always a factor of P_i, and since $-n<\nu<-n+1$, only this factor is positive, leaving behind $2n-2$ negative factors. Therefore each $P_i>0$ in this interval.

Outline of proof for Lemma (2b). After verification for $m=2 \leqq n$, induct on m via the recurrence,

$$(\nu+1)\sigma_m(\nu) = \sum_{k=1}^{m-1} \sigma_k(\nu+1)\sigma_{m-k}(\nu) \quad \text{(see [1, (20)])}.$$

which may be written as

$$(\nu+1)\sigma_m(\nu)=\sigma_1(\nu+1)\sigma_{m-1}(\nu)+\sum_{k=1}^{m-2}\{\sigma_{k+1}(\nu+1)/\sigma_k(\nu+1)\}\sigma_k(\nu+1)\sigma_{m-1-k}(\nu).$$

Here $k\leqq m-2$, so that the quantity in the braces satisfies the inductive hypothesis. Thus

$$\begin{aligned}(\nu+1)\sigma_m(\nu)&<\sigma_1(\nu+1)\sigma_{m-1}(\nu)+\frac{m-1}{2n(n+4)}\sum_{k=1}^{m-2}\sigma_k(\nu+1)\sigma_{m-1-k}(\nu)\\&=\sigma_1(\nu+1)\sigma_{m-1}(\nu)+\frac{m-1}{2n(n+4)}(\nu+1)\sigma_{m-1}(\nu).\end{aligned}$$

Divide both sides by $(\nu+1)\sigma_{m-1}(\nu)$ and use $\nu\leqq -n-2$ to get

$$\frac{\sigma_m(\nu)}{\sigma_{m-1}(\nu)}<\frac{1}{4n(n+1)}+\frac{m}{2n(n+4)}-\frac{1}{2n(n+4)}<\frac{m}{2n(n+4)}.$$

Outline of proof for Lemma 2(c). Write the recurrence formula,

$$(\nu+n)\sigma_n(\nu)=\sum_{k=1}^{n-1}\sigma_k(\nu)\sigma_{n-k}(\nu) \tag{5e}$$

as

$$(\nu+n)\sigma_n(\nu)=2\sigma_1(\nu)\sigma_{n-1}(\nu)+\sum_{k=1}^{n-3}\sigma_{k+1}(\nu)\sigma_{n-1-k}(\nu). \tag{5f}$$

Here apply the formula (5e) to $\sigma_{n-1}(\nu)$, so that (5f) becomes

$$(\nu+n)\sigma_n(\nu)=2\sigma_1(\nu)\frac{1}{\nu+n-1}\sum_{k=1}^{n-2}\sigma_k(\nu)\sigma_{n-1-k}(\nu)+\sum_{k=1}^{n-3}\sigma_{k+1}(\nu)\sigma_{n-1-k}(\nu).$$

Multiply both sides by $\nu+n-1$ to get

$$\begin{aligned}(\nu+n-1)(\nu+n)\sigma_n(\nu)&=2\sigma_1(\nu)\sum_{k=1}^{n-2}\sigma_k(\nu)\sigma_{n-1-k}(\nu)\\&\quad+(\nu+n-1)\sum_{k=1}^{n-3}\sigma_{k+1}(\nu)\sigma_{n-1-k}(\nu).\end{aligned} \tag{5g}$$

Here break up each sum $\sum_{k=1}^{n-\cdot}$ into two sums $\sum_{k=1}^{n-5}+\sum_{k=n-4}^{n-}$, so that the right side of (5g) becomes, after we collect terms and expand,

$$\begin{aligned}&\left[\sum_{k=1}^{n-5}\{(\nu+n-1)\sigma_{k+1}(\nu)+2\sigma_1(\nu)\sigma_k(\nu)\}\sigma_{n-1-k}(\nu)\right]+2\sigma_1(\nu)\sigma_3(\nu)\sigma_{n-4}(\nu)\\&\quad+\left[\sum_{k=1}^{2}\{(\nu+n-1)\sigma_{k+1}(\nu)+2\sigma_1(\nu)\sigma_k(\nu)\}\sigma_{n-1-k}(\nu)\right].\end{aligned} \tag{5h}$$

Divide inside and multiply outside of each pair of the braces, $\{\cdot\}$, by $\sigma_k(\nu)$, to introduce the new quantity,

$$(\nu+n-1)\frac{\sigma_{k+1}(\nu)}{\sigma_k(\nu)}+2\sigma_1(\nu), \tag{†}$$

to which, since $k+1 \leqq n$, Lemma 2(b) applies; then (†) reduces to

$$< \frac{\nu+n-1}{2(n+4)} + \frac{1}{2(\nu+1)} = \frac{\nu+n-1}{2(n+4)} - \frac{1}{(-2)(\nu+1)}$$

$$< \frac{\nu+n-1}{2(n+4)} - \frac{1}{2(n-2)} < 0, \quad \text{for } -n+1<\nu<-n+2.$$

Thus $(\dagger)<0$, on $(-n+1, -n+2)$.

Since $k \leqq n-5$, and $\nu<-n+2$, then $\nu<-k$. Therefore $\sigma_k(\nu)<0$ [2, (18)]. Similarly each of $\sigma_{n-1-k}(\nu)$, $\sigma_1(\nu)$, $\sigma_3(\nu)$, and $\sigma_{n-4}(\nu)$ is <0 in (5h). This result, together with $(\dagger)<0$, implies that the quantity in (5h) (and hence (5g)) is negative. Thus,

$$(\nu+n-1)(\nu+n)\sigma_n(\nu)<0 \quad \text{for } -n+1<\nu<-n+2.$$

Outline of proof for Lemma 2(d). The reader may check from the functions listed at end of article that except for $n=5$, $\sigma_n(\nu)>0$ everywhere in $(-n+2, -n+3)$, for $1 \leqq n \leqq 8$. We therefore assume below that $n>8$. Write (5e) as

$$\text{(5i)} \qquad (\nu+n)\sigma_n(\nu) = 2\sigma_1(\nu)\sigma_{n-1}(\nu) + 2\sigma_2(\nu)\sigma_{n-2}(\nu) + \sum_{k=1}^{n-3} \sigma_k(\nu)\sigma_{n-k}(\nu).$$

Now apply formula (5e) to $\sigma_{n-1}(\nu)$ and $\sigma_{n-2}(\nu)$ above; also write the quantity on the extreme right of the above equation as the sum of the first two terms and the rest; collect terms, readjust subscripts so that the above is equal to

$$\frac{2\sigma_1(\nu)}{\nu+n-1}\sum_{k=1}^{n-2} \sigma_k(\nu)\sigma_{n-1-k}(\nu) + \frac{2\sigma_2(\nu)}{\nu+n-2}\sum_{k=1}^{n-3} \sigma_k(\nu)\sigma_{n-2-k}(\nu)$$

$$+\sigma_1\sigma_{n-1}(\nu) + \sigma_2\sigma_{n-2}(\nu) + \sum_{k=1}^{n-5} \sigma_{k+2}\sigma_{n-2-k}(\nu).$$

First apply formula (5e) again to σ_{n-1} and σ_{n-2} and collect terms. But

$$\sum_{k=1}^{n-2} \sigma_k\sigma_{n-1-k}(\nu) = 2\sigma_1\sigma_{n-2}(\nu) + \sum_{k=2}^{n-3} \sigma_k\sigma_{n-1-k}(\nu)$$

so the above reduces to

$$\frac{6\sigma_1^2(\nu)}{\nu+n-1}\cdot\sigma_{n-2}(\nu) + \frac{3\sigma_1(\nu)}{\nu+n-1}\sum_{k=2}^{n-3} \sigma_k\sigma_{n-1-k}(\nu)$$

$$+\frac{3\sigma_2(\nu)}{\nu+n-2}\sum_{k=1}^{n-3} \sigma_k\sigma_{n-2-k}(\nu)$$

$$+\sum_{k=1}^{n-5} \sigma_{k+2}\sigma_{n-2-k}(\nu).$$

Another use of formula (5e) on σ_{n-2} and a readjustment of subscripts in the second summation above reduce the above to

$$\frac{6\sigma_1^2(\nu)}{(\nu+n-1)(\nu+n-2)}\sum_{k=1}^{n-3} \sigma_k\sigma_{n-2-k}(\nu) + \frac{3\sigma_1(\nu)}{\nu+n-1}\sum_{k=1}^{n-4} \sigma_{k+1}\sigma_{n-2-k}(\nu)$$

$$+\frac{3\sigma_2(\nu)}{\nu+n-2}\sum_{k=1}^{n-3} \sigma_k\sigma_{n-2-k}(\nu) + \sum_{k=1}^{n-5} \sigma_{k+2}\sigma_{n-2-k}(\nu).$$

Multiply this result, as well as the left side of (5i), by $(\nu+n-2)(\nu+n-1)$. After canceling out redundant factors, break up each summation (except the last one) into two sub-summations so as to get

$$(5j)\quad (\nu+n-2)(\nu+n-1)(\nu+n)\sigma_n(\nu)$$

$$= \left[6\sigma_1^2(\nu)\sum_{k=1}^{n-6}\sigma_k\sigma_{n-2-k}(\nu)\right]+\left[6\sigma_1^2(\nu)\sum_{k=n-5}^{n-3}\sigma_k\sigma_{n-2-k}(\nu)\right]$$

$$+\left[3(\nu+n-2)\sigma_1(\nu)\sum_{k=1}^{n-6}\sigma_{k+1}(\nu)\sigma_{n-2-k}(\nu)\right]$$

$$+\left[3(\nu+n-2)\sigma_1(\nu)\sum_{k=n-5}^{n-4}\sigma_{k+1}\sigma_{n-2-k}(\nu)\right]$$

$$+\left[3(\nu+n-1)\sigma_2(\nu)\sum_{k=1}^{n-6}\sigma_k\sigma_{n-2-k}(\nu)\right]$$

$$+\left[3(\nu+n-1)\sigma_2(\nu)\sum_{k=n-5}^{n-3}\sigma_k\sigma_{n-2-k}(\nu)\right]$$

$$+\left[(\nu+n-1)(\nu+n-2)\sum_{k=1}^{n-5}\sigma_{k+2}\sigma_{n-2-k}(\nu)\right]$$

$$=[6\sigma_1^2(\nu)+3(\nu+n-1)\sigma_2(\nu)]\sum_{k=1}^{n-6}\sigma_k\sigma_{n-2-k}(\nu)$$

$$+\left[3(\nu+n-2)\sigma_1(\nu)\sum_{k=1}^{n-6}\sigma_{k+1}\sigma_{n-2-k}(\nu)\right]$$

$$+[6\sigma_1^2(\nu)+3(\nu+n-1)\sigma_2(\nu)]\sum_{k=n-5}^{n-3}\sigma_k\sigma_{n-2-k}(\nu)$$

$$+\left[3(\nu+n-2)\sigma_1(\nu)\sum_{k=n-5}^{n-4}\sigma_{k+1}\sigma_{n-2-k}(\nu)\right]$$

$$+(\nu+n-1)(\nu+n-2)\sum_{k=1}^{n-5}\sigma_{k+2}\sigma_{n-2-k}(\nu)$$

$$(5k)\quad =\sum_{k=1}^{n-6}\left\{3(\nu+n-2)\sigma_1(\nu)\frac{\sigma_{k+1}(\nu)}{\sigma_k(\nu)}+3(\nu+n-1)\sigma_2(\nu)+6\sigma_1^2(\nu)\right\}\sigma_k\sigma_{n-2-k}(\nu)$$

$$(5l)\quad +\sum_{k=1}^{2}\left\{3(\nu+n-2)\sigma_1(\nu)\frac{\sigma_{k+1}(\nu)}{\sigma_k(\nu)}+3(\nu+n-1)\sigma_2(\nu)+6\sigma_1^2(\nu)\right\}\sigma_k\sigma_{n-2-k}(\nu)$$

$$(5m)\quad +\{3(\nu+n-1)\sigma_2(\nu)+6\sigma_1^2(\nu)\}\sigma_3(\nu)\sigma_{n-5}(\nu)$$

$$(5n)\quad +(\nu+n-1)(\nu+n-2)\sum_{k=1}^{n-5}\sigma_{k+2}(\nu)\sigma_{n-2-k}(\nu).$$

In (5k), $k\leqq n-6$, so that $k+1\leqq n-5$; then since $-n+2<\nu<-n+3$, we have $\nu<-(k+1)$, $\nu<-k$, and $\nu<-(n-2-k)$. Therefore [2(18)] $\sigma_k(\nu)$, $\sigma_{k+1}(\nu)$, $\sigma_{n-2-k}(\nu)$ are all negative in (5k). It is now easy to see that (5k)>0. Similarly, all of (5l)–(5n) are positive on $-n+2<\nu<-n+3$. Thus the left side of (5j) is positive in this interval and the result follows.

THEOREM. *Let $n>5$. If $-n+1$, $-n+2$, and $-n+3$ are not roots of $\phi_n(\nu)$, then every real root lies in $(-n+3, -2)$, and $\phi_n>0$ on $(-\infty, -n+3]$ if $d(n)$ is even, and $\phi_n<0$ otherwise.*

Proof. By (4c) every real root lies in $[-n+1, -2)$ so we only consider that interval. Let

$$r=\sum_{k=1}^{n-2}[n/k]; \qquad t=\sum_{k=1}^{n-3}[n/k]; \qquad \psi_1\equiv\prod_{s=1}^{r}(\nu+b_s), \qquad 1\leqq b_s\leqq n-2;$$

$$\psi_2\equiv\prod_{s=1}^{2}(\nu+a_s), \qquad n-1\leqq a_s\leqq n; \qquad \psi_3\equiv\prod_{s=1}^{t}(\nu+b_s'), \qquad 1\leqq b_s'\leqq n-3;$$

$$\psi_4\equiv\prod_{s=1}^{3}(\nu+a_s'), \qquad n-2\leqq a_s'\leqq n.$$

Then (1a) can be written in the form

$$\phi_n(\nu)=4^n\psi_1\psi_2\sigma_n(\nu)=4^n\psi_3\psi_4\sigma_n(\nu).$$

For $n\geqq 5$,

$$\left[\frac{n}{n-2}\right]=\left[\frac{n}{n-1}\right]=\left[\frac{n}{n}\right]=1,$$

so in view of (2ii) $r=d(n)+2n-3$ and $t=d(n)+2n-4$. Now if $d(n)$ is even, then r is odd and t is even, so on $(-n+1, -n+2)$ we get $\psi_1<0$, $\psi_2>0$, while on $(-n+2, -n+3)$ we get $\psi_3>0$, $\psi_4>0$. But by Lemma 2, $\sigma_n(\nu)<0$ on $(-n+1, -n+2)$ and >0 on $(-n+2, -n+3)$. Therefore,

(5p) $\phi_n(\nu)>0$ on both $(-n+1, -n+2)$ and $(-n+2, -n+3)$ if $d(n)$ is even.

Similarly,

(5q) $\phi_n(\nu)<0$ on both $(-n+1, -n+2)$ and $(-n+2, -n+3)$ if $d(n)$ is odd.

The result follows.

Remark. The real nonintegral zeros of ϕ_n are seen therefore to lie within $(-n+3, -2)$, $n>5$. It is believed that when $n>8$ all these zeros lie probably within only $(-[n/2]-1, -2)$, and that, in any case, at least one of them lies in $(-3, -2)$. The second part of the preceding statement is true for at least the first eight ϕ_n's. Furthermore, it is not known whether there is an n for which ϕ_n has an integral zero. A conjecture of Kishore [3, p. 518] states that there is not.

3. The polynomial on $[-2, \infty)$. ϕ_n behaves here nearly the same way as it does on $(-\infty, -n+3)$, except that now $\phi_n(\nu)>0$ for all n. We show directly from (2) that $\phi_n(\nu)\geqq 1$ on $\nu\geqq -2$, for all n; in fact that on $\nu\geqq -1$, $\phi_n(\nu)\geqq e_n$. We also obtain a dominating polynomial.

By induction on n, with a recurrence formula for $\phi_n(\nu)$, it is shown in [2, § 5] that $\phi_n>0$ on $\nu\geqq -2$. Then in view of (2), there is k, $1\leqq k\leqq c(n)$, such that the component

$$2^{n_k}\prod_{j=2}^{n-1}(\nu+j)^{n_{kj}},$$

of ϕ_n is positive at $\nu=-2$. Then $\nu+2$ cannot be a factor of this component ($n_{k2}=0$), and we write this component in the form

$$2^{n_k}\prod_{j=3}^{n-1}(\nu+j)^{n_{kj}}.$$

Thus, with $j \geqq 3$ and each component of (2) nonnegative on $[-2, \infty)$, we obtain

$$\phi_n(-2) \geqq 2^{n_k}\prod_{j=3}^{n-1}(-2+j)^{n_{kj}} \geqq 1. \tag{6}$$

In (3) it is clear that $\phi_n' > 0$ on $\nu > -2$, so that ϕ_n is strictly increasing on $(-2, \infty)$. By continuity it follows also that ϕ_n is strictly increasing on all of $[-2, \infty)$. This and (6) lead to

$$\phi_n(\nu) \geqq 1 \quad \text{on } [-2, \infty). \tag{6a}$$

Next fix ε, $0 \leqq \varepsilon \leqq 1$. Then on $[-(1+\varepsilon), \infty)$, $\nu + b_{is} \geqq 1-\varepsilon > 0$ for each b_{is}; and hence, from (2a),

$$\phi_n(\nu) \geqq \frac{1}{n}\binom{2n-2}{n-1}(1-\varepsilon)^{d(n)} \quad \text{on } \nu \geqq -(1+\varepsilon), \tag{6b}$$

for all n, with equality holding when $n = 1, 2, 3$. Letting $\varepsilon \to 0$, we see that

$$\phi_n(\nu) \geqq \frac{1}{n}\binom{2n-2}{n-1} \quad \text{on } [-1, \infty). \tag{6c}$$

As before, using Lemma 1, we see that

$$\phi_n'' > 0 \quad \text{on } [-(1+\varepsilon), \infty), \quad \text{for each } n > 5; \tag{6d}$$

while by using (2a), we see that

$$\phi_n(\nu) \leqq \sum_{i=1}^{c} 2^{n_i}\prod_{s=1}^{d(n)}(\nu+n-1) = \frac{1}{n}\binom{2n-2}{n-1}(\nu+n-1)^{d(n)} \equiv h_n(\nu). \tag{6e}$$

Summarizing the foregoing, we have:
(6f) The Rayleigh polynomial of order n lies above the lines

$$y = 1, \quad y = \frac{1}{n}\binom{2n-2}{n-1} \quad \text{and} \quad y = h_n(-n+2-\varepsilon)$$

on the intervals $\nu \geqq -2$, $\nu \geqq -1$ and $\nu \geqq -(1+\varepsilon)$ respectively; $\phi_n(\nu)$ is strictly increasing $(n > 3)$ on $\nu \geqq 2$, concave upwards $(n > 5)$ on $\nu \geqq -2$, and majorized by the polynomial $h_n(\nu)$ on $[-2, \infty)$ for all $n \geqq 1$.

We remark in passing that on $[-n+1, -2]$, $\phi_n(\nu)$ lies too far within the horizontal strip $y = \pm h_n(-2)$ for $n > 3$, so that there, $|\phi_n| \ll h_n(-2)$; and that in view of (4a), $|\phi_n| \leqq h_n(-\nu-n-1)$ on $\nu \leqq -n+1$ by continuity, for $n > 3$; and that by expanding (2a) the leading coefficient of $\phi_n(\nu)$ is

$$\sum 2^{n_i} = \frac{1}{n}\binom{2n-2}{n-1} \qquad \text{(see [3, (17)])},$$

which equals the leading coefficient of $h_n(\nu)$, so that since the two polynomials have the same degrees,

$$\lim_{\nu\to\infty}\frac{\phi_n(\nu)}{h_n(\nu)} = 1.$$

Also in view of (6a), $\sum_{n=1}^{\infty} \phi_n(\nu)$ is a divergent series at each point of at least the region $[-2, \infty)$. This is in contrast with $\sum_{n=1}^{\infty} \sigma_n(\nu)$, which converges uniformly (see [7, § 4]) on $[0, \infty)$.

We have seen in §§ 2 and 3 some roles played by the Catalan numbers, e_n, in ϕ_n. These numbers appear quite often in combinatorial mathematics. It is known that if P_n is the family of all the ways in which a given planar regular $(n+1)$gon can be partitioned into triangles by $n-2$ diagonals that do not intersect inside the polygon, B_n is the family of all binary products of a given element taken n times, where these products are assumed to be neither commutative nor associative, and if S_n is the set of all the points, $x=(x_i)_{i=1}^{2n-2} \in R^{2n-2}$, such that $x_1 = \pm 1$ and $x_1 + \cdots + x_k \geqq 0$ if $k < 2n+2$, and 0 if $k = 2n+2$, then all the above sets have the same cardinality: their common cardinal number is e_n. As we already know, the leading coefficient of every Rayleigh polynomial of order n is e_n, while by § 3, $\phi_n(\nu)$ lies between e_n and $e_n(\nu+n-1)^{d(n)}$ on $[-1, \infty)$. L. Shapiro [in Amer. Math. Monthly, 82 (1975), no. 6, p. 634] has shown that the number of ideals in the ring of $n \times n$ upper triangular matrices is e_{n-2}. Finally, for the nth order Rayleigh function, it can be shown via (5) and (5d) that $\sigma_n(\nu)$ on $(-1, \infty)$ lies between $e_n 4^{-n}(\nu+n)^{1-2n}$ and $e_n 4^{-n}(\nu+1)^{1-2n}$. The first eight polynomials are listed below for reference purposes, as numerators of corresponding $\sigma_n(\nu)$ (see [2, (2)]):

$$1.\ \frac{1}{4(\nu+1)};\quad 2.\ \frac{1}{4^2(\nu+1)^2(\nu+2)};\quad 3.\ \frac{2}{4^3(\nu+1)^3(\nu+2)(\nu+3)};$$

$$4.\ \frac{5\nu+11}{4^4(\nu+1)^4(\nu+2)^2(\nu+3)(\nu+4)};\quad 5.\ \frac{14\nu+38}{4^5(\nu+1)^5(\nu+2)^2(\nu+3)(\nu+4)(\nu+5)};$$

$$6.\ \frac{42\nu^3+362\nu^2+1026\nu+946}{4^6(\nu+1)^6(\nu+2)^3(\nu+3)^2(\nu+4)(\nu+5)(\nu+6)};$$

$$7.\ \frac{132\nu^3+1316\nu^2+4324\nu+3580}{4^7(\nu+1)^7(\nu+2)^3(\nu+3)^2(\nu+4)(\nu+5)(\nu+6)};$$

$$8.\ \frac{429\nu^5+7640\nu^4+53752\nu^3+185430\nu^2+311387\nu+202738}{4^8(\nu+1)^8(\nu+2)^4(\nu+3)^2(\nu+4)^2(\nu+5)(\nu+6)(\nu+7)(\nu+8)}.$$

REFERENCES

[1] NAND KISHORE, *The Rayleigh function*, Proc. Amer. Math. Soc., 14 (1963) no. 4, pp. 527–533.
[2] ———, *The Rayleigh polynomial*, Ibid., 15(1964), no. 6. Vol. 15, pp. 911–917.
[3] ———, *A structure of the Rayleigh polynomial*, Duke Math. J., 31 (1964), no. 3, pp. 513–518.
[4] ———, *A representation of Bernoulli numbers*, Pacific J. Math., 14 (1964), no. 4, pp. 1297–1304.
[5] ———, *A class of formulas for the Rayleigh function*, Duke Math. J., 34 (1967), no. 3, pp. 573–580.
[6] ———, Lecture notes, Univ. of Toledo, Ohio, 1969.
[7] E. C. OBI, *Functional bounds, series and log convexity of Rayleigh higher derivatives*; J. Math. Anal. Appl., 52 (1975), no. 3, pp. 648–659.
[8] G. N. WATSON, *A Treatise on the Theory of Bessel Functions*, Macmillan, New York, 1944.

SIAM J MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0007 $01.0070

ASYMPTOTIC EXPANSIONS OF FRACTIONAL INTEGRALS INVOLVING LOGARITHMS*

R. WONG†

Abstract. Let $\phi(t)$ be a locally integrable function on $[0, \infty)$ and satisfy

$$\phi(t) \sim t^{-\beta} \sum_{m=0}^{\infty} c_m (\ln t)^{\gamma-m} \qquad \text{as } t \to \infty,$$

where $\beta \geqq 0$ and γ is arbitrary. Asymptotic expansions are obtained for the fractional integral of order α defined by

$$I^\alpha \phi(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \phi(s)\, ds, \qquad \alpha > 0.$$

1. Introduction. Let ϕ be a locally integrable function on $[0, \infty)$ and let $\operatorname{Re} \alpha > 0$. The operator I^α of integration of order α is defined by

$$I^\alpha \phi(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \phi(s)\, ds. \tag{1.1}$$

In a recent paper [1], Berger and Handelsman have studied the asymptotic behavior of $I^\alpha \phi(t)$ as $t \to \infty$, when $\phi(t)$ satisfies

$$\phi(t) \sim \exp(-ct^p) \sum_{m=0}^{\infty} d_m t^{-r_m} \qquad \text{as } t \to \infty, \tag{1.2}$$

where $p > 0$, $\operatorname{Re} c \geqq 0$ and $\operatorname{Re} r_m \uparrow \infty$ as $m \to \infty$.

In this paper we consider the case

$$\phi(t) \sim t^{-\beta} \sum_{m=0}^{\infty} c_m (\ln t)^{\gamma-m} \qquad \text{as } t \to \infty \tag{1.3}$$

where $\operatorname{Re} \beta \geqq 0$ and γ is arbitrary. Functions having the asymptotic form (1.3) arise in various problems of applied mathematics [3], [6].

Our investigation is motivated by a study of the nonlinear integral equation

$$\phi(t) = \frac{1}{\sqrt{\pi}} \int_0^t (t-s)^{-1/2} \{f(s) - \phi^n(s)\}\, ds, \qquad n \geqq 1, \tag{1.4}$$

where $f(t)$ is nonnegative, bounded and locally integrable on $[0, \infty)$ and has an asymptotic expansion of the form

$$f(t) \sim \sum_{m=0}^{\infty} \gamma_m t^{-a_m} \qquad \text{as } t \to \infty, \tag{1.5}$$

with $\gamma_0 > 0$ and $a_0 < a_1 < a_2 < \cdots$. In [4] and [5], Handelsman and Olmstead have obtained the dominant term (and in many cases the leading two terms) of the formal asymptotic solution to (1.4) for various values of n and a_0. We are particularly

* Received by the editors April 14, 1976, and in final revised form February 2, 1977.

† Department of Mathematics, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2. This research was partially supported by the National Research Council of Canada under Contract A7359.

interested in the special case $n=2$ and $a_0>1$ since the solution $\phi(t)$ has the asymptotic behavior

$$\phi(t) \sim \frac{\sqrt{\pi}}{\sqrt{t}\ln t} \qquad \text{as } t\to\infty, \tag{1.6}$$

which is of the form being considered in the present paper.

Recently Bleistein [2] has developed a technique for obtaining the asymptotic expansion of a class of integral transforms of functions whose asymptotic expansions near the origin involve arbitrary powers of ln t. Since the fractional integral (1.1) may be put in a form similar to that considered by Bleistein, his results are closely related to ours, although the two methods are quite different.

2. Asymptotic behavior of $I^\alpha\phi(t)$. For the sake of simplicity it will be assumed throughout the paper that the parameters α, β and γ are real, and that the function $\phi(t)$ is real and nonnegative. The extension to complex values of the parameters is automatic and the extension to allow complex-valued functions ϕ will not present any difficulty. Also for simplicity we divide the discussion into three cases: (i) $0\leqq\beta<1$, (ii) $\beta=1$ and (iii) $\beta>1$. However, only Cases (i) and (ii) will be considered in detail. We omit the third case, since it is very similar to the first two.

Case (i). Returning to (1.1), we write

$$I^\alpha\phi(t)=I_1^\alpha\phi(t)+I_2^\alpha\phi(t), \tag{2.1}$$

where $I_1^\alpha\phi(t)$ and $I_2^\alpha\phi(t)$ correspond, respectively, to the intervals $(0,\sqrt{t})$ and $(\sqrt{t}, t)$.

LEMMA 1. *If $\phi(t)$ satisfies* (1.3) *with $0\leqq\beta<1$ then there exists a fixed $\rho>0$ such that*

$$I_1^\alpha\phi(t)=O(t^{\alpha-\beta-\rho}) \qquad \textit{as } t\to\infty. \tag{2.2}$$

Proof. Choose $0<\varepsilon<1-\beta$. Since $(\ln t)^\gamma=O(t^\varepsilon)$ as $t\to\infty$, it follows from (1.3) that $\phi(t)=O(t^{-\beta+\varepsilon})$ as $t\to\infty$ and also that

$$\int_0^{\sqrt{t}}\phi(s)\,ds=O(t^{(1-\beta+\varepsilon)/2}) \qquad \text{as } t\to\infty. \tag{2.3}$$

Set $M_\alpha=\max_{0\leqq u\leqq 1/2}(1-u)^{\alpha-1}$. Then $(t-s)^{\alpha-1}\leqq M_\alpha t^{\alpha-1}$ for $0\leqq s\leqq\sqrt{t}$ and

$$\Gamma(\alpha)I_1^\alpha\phi(t)\leqq M_\alpha t^{\alpha-1}\int_0^{\sqrt{t}}\phi(s)\,ds. \tag{2.4}$$

Coupling (2.3) and (2.4), we obtain the result (2.2) with $\rho=\frac{1}{2}(1-\beta-\varepsilon)$.

LEMMA 2. *For $\alpha>0$, $0\leqq\beta<1$ and any integer $k\geqq0$, there exists a fixed $\delta>0$ such that*

$$\int_0^{1/\sqrt{t}}(1-u)^{\alpha-1}u^{-\beta}(\ln u)^k\,du=O(t^{-\delta}) \qquad \textit{as } t\to\infty. \tag{2.5}$$

Proof. We again choose $0<\varepsilon<1-\beta$ and observe that $(\ln u)^k=O(u^{-\varepsilon})$ as $u\to0^+$. Hence the integral in (2.5) is dominated by

$$M_\alpha\int_0^{1/\sqrt{t}}u^{-\beta-\varepsilon}\,du=O(t^{-(1-\beta-\varepsilon)/2}).$$

The result (2.5) now follows by letting $\delta=\frac{1}{2}(1-\beta-\varepsilon)$.

In order to simplify some of the expressions which occur, we introduce the following notations:

$$L_\alpha(\beta, \gamma; t) = \int_{\sqrt{t}}^{t} (t-s)^{\alpha-1} s^{-\beta} (\ln s)^\gamma \, ds \tag{2.6}$$

$$\Omega_k(\alpha, \beta) = \int_0^1 (1-u)^{\alpha-1} u^{-\beta} (\ln u)^k \, du. \tag{2.7}$$

In terms of the Beta function, we have

$$\Omega_k(\alpha, \beta) = (-1)^k \frac{d^k}{d\beta^k} B(\alpha, 1-\beta). \tag{2.8}$$

LEMMA 3. *For $\alpha > 0$, $0 \leqq \beta < 1$ and γ arbitrary,*

$$L_\alpha(\beta, \gamma; t) \sim t^{\alpha-\beta} \sum_{k=0}^{\infty} \binom{\gamma}{k} \Omega_k(\alpha, \beta)(\ln t)^{\gamma-k}, \qquad \textit{as } t \to \infty. \tag{2.9}$$

Proof. In (2.6) we make the substitution $s = ut$. Then

$$\begin{aligned} L_\alpha(\beta, \gamma; t) &= t^{\alpha-\beta} \int_{1|\sqrt{t}}^{1} (1-u)^{\alpha-1} u^{-\beta} (\ln t + \ln u)^\gamma \, du \\ &= t^{\alpha-\beta} (\ln t)^\gamma \int_{1/\sqrt{t}}^{1} (1-u)^{\alpha-1} u^{-\beta} \left(1 + \frac{\ln u}{\ln t}\right)^\gamma du. \end{aligned} \tag{2.10}$$

On the path of integration $t^{-1/2} \leqq u \leqq 1$, we have $|(\ln u)/\ln t| \leqq \frac{1}{2}$. Hence, for every fixed integer $K \geqq 0$,

$$\left(1 + \frac{\ln u}{\ln t}\right)^\gamma = \sum_{k=0}^{K} \binom{\gamma}{k} \frac{(\ln u)^k}{(\ln t)^k} + O\left(\frac{(\ln u)^{K+1}}{(\ln t)^{K+1}}\right). \tag{2.11}$$

Inserting (2.11) into(2.10) and carrying out the integration term by term, we obtain from (2.7) and Lemma 2

$$L_\alpha(\beta, \gamma; t) = t^{\alpha-\beta} \sum_{k=0}^{K} \binom{\gamma}{k} \Omega_k(\alpha, \beta)(\ln t)^{\gamma-k} + R_K(t).$$

Since the integral

$$\int_{1/\sqrt{t}}^{1} (1-u)^{\alpha-1} u^{-\beta} (\ln u)^{K+1} \, du$$

exists and is bounded as $t \to \infty$, the remainder term $R_K(t)$ satisfies

$$R_K(t) = O(t^{\alpha-\beta} (\ln t)^{\gamma-K-1}) \qquad \text{as } t \to \infty. \tag{2.12}$$

The O-terms from (2.5) are included in that appearing in (2.12), and this lemma is thus proved.

With the aid of these three preliminary results, we are now ready to state and prove our main theorem.

THEOREM 1. *Let ϕ be locally integrable on* $[0, \infty)$ *and satisfy* (1.3) *with* $0 \leqq \beta < 1$. *Then as $t \to \infty$,*

$$I^\alpha \phi(t) \sim t^{\alpha-\beta} \sum_{l=0}^{\infty} b_l (\ln t)^{\gamma-l}, \tag{2.13}$$

where the coefficients b_l are constants explicitly given by

$$b_l=\frac{1}{\Gamma(\alpha)}\sum_{k=0}^{l} c_{l-k}\binom{\gamma+k-l}{k}\Omega_k(\alpha,\beta). \tag{2.14}$$

Proof. From (2.1) and Lemma 1, we have

$$I^\alpha\phi(t)=I_2^\alpha\phi(t)+O(t^{\alpha-\beta-\rho}) \quad \text{as } t\to\infty, \tag{2.15}$$

where ρ is a fixed positive number. Writing

$$\phi(t)=\sum_{m=0}^{M} c_m t^{-\beta}(\ln t)^{\gamma-m}+R_M(t)$$

gives

$$I_2^\alpha\phi(t)=\frac{1}{\Gamma(\alpha)}\sum_{m=0}^{M} c_m L_\alpha(\beta,\gamma-m;t)+r_M(t) \tag{2.16}$$

where

$$r_M(t)=\frac{1}{\Gamma(\alpha)}\int_{\sqrt{t}}^{t}(t-s)^{\alpha-1}R_M(s)\,ds.$$

From (1.3), it follows that there are constants $K>0$ and $c>1$ such that

$$|R_M(t)|\leqq Kt^{-\beta}(\ln t)^{\gamma-M-1} \quad \text{for } t\geqq c.$$

Hence, by Lemma 3,

$$\begin{aligned} |r_M(t)| &< \frac{K}{\Gamma(\alpha)}\int_{\sqrt{t}}^{t}(t-s)^{\alpha-1}s^{-\beta}(\ln s)^{\gamma-M-1}\,ds \\ &= O(t^{\alpha-\beta}(\ln t)^{\gamma-M-1}) \quad \text{as } t\to\infty. \end{aligned} \tag{2.17}$$

Combining the results (2.16) and (2.17), we obtain

$$I_2^\alpha\phi(t)=\frac{1}{\Gamma(\alpha)}\sum_{m=0}^{M} c_m L_\alpha(\beta,\gamma-m;t)+O(t^{\alpha-\beta}(\ln t)^{\gamma-M-1}) \tag{2.18}$$

as $t\to\infty$, and hence

$$I^\alpha\phi(t)=\frac{1}{\Gamma(\alpha)}\sum_{m=0}^{M} c_m L_\alpha(\beta,\gamma-m;t)+O(t^{\alpha-\beta}(\ln t)^{\gamma-M-1}) \tag{2.19}$$

as $t\to\infty$, since the O-term appearing in (2.15) may be included in that appearing in (2.18). Substituting (2.9) in (2.19) and regrouping the terms, we have for any $L\geqq 0$

$$I^\alpha\phi(t)=t^{\alpha-\beta}\left[\sum_{l=0}^{L} b_l(\ln t)^{\gamma-l}+O((\ln t)^{\gamma-L-1})\right]$$

as $t\to\infty$, where the coefficients b_l are given in (2.14). This completes the proof of Theorem 1.

Case (ii). In place of (2.1) we write in this case

$$I^\alpha\phi(t)=\frac{1}{\Gamma(\alpha)}t^{\alpha-1}\int_0^t\phi(s)\,ds+\frac{1}{\Gamma(\alpha)}\int_0^t[(t-s)^{\alpha-1}-t^{\alpha-1}]\phi(s)\,ds, \tag{2.20}$$

and work only with the second integral on the right. Divide the range of integration at $s=\sqrt{t}$, and denote by $J_1^\alpha\phi(t)$ and $J_2^\alpha\phi(t)$ the integrals corresponding, respectively, to the intervals $(0,\sqrt{t})$ and $(\sqrt{t},t)$. In analogy to (2.2) we have an estimate for $J_1^\alpha\phi(t)$, which is asymptotically negligible with respect to $J_2^\alpha\phi(t)$. Proceeding as in Theorem 1 and Lemma 3, we then obtain an asymptotic series for $J_2^\alpha\phi(t)$ in powers of $(\ln t)^{-1}$. Since the argument here is similar to that for Case (i) we omit the details. The final result is stated in the following theorem.

THEOREM 2. *Let $\phi(t)$ be locally integrable on* $[0, \infty)$ *and satisfy* (1.3) *with* $\beta = 1$. *Put*

$$\Lambda_k(\alpha) = \int_0^1 [(1-u)^{\alpha-1} - 1]u^{-1}(\ln u)^k du. \tag{2.21}$$

Then as $t \to \infty$

$$t^{1-\alpha} I^\alpha\phi(t) \sim \frac{1}{\Gamma(\alpha)} \int_0^1 \phi(s)\, ds + \sum_{l=0}^{\infty} d_l (\ln t)^{\gamma-l}, \tag{2.22}$$

where the coefficients d_l are constants given explicitly by

$$d_l = \frac{1}{\Gamma(\alpha)} \sum_{k=0}^{l} c_{l-k} \binom{\gamma+k-l}{k} \Lambda_k(\alpha). \tag{2.23}$$

A simple calculation gives

$$\Lambda_0(\alpha) = \psi(1) - \psi(\alpha), \tag{2.24}$$

where $\psi(\alpha) = \Gamma'(\alpha)/\Gamma(\alpha)$.

Remark. At the beginning of the section, we remarked that Case (iii), in which $\beta > 1$ in (1.3), is similar to Cases (i) and (ii). From the above analysis it should now become clear that we can indeed give the asymptotic behavior of $I^\alpha\phi(t)$ in the third case. The result is, however, more complicated than (2.13) and (2.22). We shall not give it explicitly.

3. Expansions involving ln ln *t*. In § 4 we shall see that the formal solution to equation (1.4) may have the asymptotic form

$$\phi(t) \sim t^{-\beta}(\ln t)^\gamma \left[c_0 + c_1 \frac{(\ln \ln t)}{(\ln t)} + c_2 \frac{(\ln \ln t)^2}{(\ln t)^2} + \cdots \right], \tag{3.1}$$

where $0 \leqq \beta \leqq 1$ and γ is arbitrary. Although the expansion is not of the type considered in (1.3), our method can again be used to give the large-t behavior of $I^\alpha\phi(t)$. The algebraic complexities, however, become overwhelming. In the following theorems, we shall content ourselves with giving the leading five terms in the asymptotic expansions.

THEOREM 3. *Let $\phi(t)$ be locally integrable on* $[0, \infty)$ *and satisfy* (3.1). *If* $0 \leqq \beta < 1$, *then as* $t \to \infty$

$$\begin{aligned} I^\alpha\phi(t) \sim t^{\alpha-\beta}(\ln t)^\gamma \bigg[& d_0 + d_1 \frac{\ln \ln t}{\ln t} + d_2 \frac{1}{\ln t} \\ & + d_3 \frac{(\ln \ln t)^2}{(\ln t)^2} + d_4 \frac{\ln \ln t}{(\ln t)^2} + \cdots \bigg], \end{aligned} \tag{3.2}$$

where $d_0 = c_0\Gamma(1-\beta)/\Gamma(\alpha-\beta+1)$, $d_1 = c_1\Gamma(1-\beta)/\Gamma(\alpha-\beta+1)$, $d_2 = c_0\gamma\Omega_1(\alpha, \beta)/\Gamma(\alpha)$, $d_3 = c_2\Gamma(1-\beta)/\Gamma(\alpha-\beta+1)$ *and* $d_4 = c_1(\gamma-1)\Omega_1(\alpha, \beta)/\Gamma(\alpha)$; $\Omega_1(\alpha, \beta)$ *being the constant given in* (2.7) (*with* $k = 1$).

Proof. For convenience, we set

$$L_{\gamma,k}(t)=(\ln t)^{\gamma-k}(\ln\ln t)^k, \qquad k=0,1,2,\cdots, \tag{3.3}$$

and rewrite (3.1) in the form

$$\phi(t)=t^{-\beta}\left[\sum_{k=0}^{K-1} c_k L_{\gamma,k}(t)+\varepsilon_K(t)\right], \tag{3.4}$$

where the remainder $\varepsilon_K(t)$ satisfies

$$\varepsilon_K(t)=O(L_{\gamma,K}(t)) \qquad \text{as } t\to\infty. \tag{3.5}$$

Let $I_1^\alpha\,\phi(t)$ and $I_2^\alpha\,\phi(t)$ be defined as in (2.1). Using the proof of Lemma 1 gives

$$I_1^\alpha\phi(t)=O(t^{\alpha-\beta-\rho}) \qquad \text{as } t\to\infty \tag{3.6}$$

for some fixed $\rho>0$. From (3.4) we have

$$I_2^\alpha\phi(t)=\frac{1}{\Gamma(\alpha)}\sum_{k=0}^{K-1} c_k N_k(\alpha,\beta,\gamma;t)+E_K(t) \tag{3.7}$$

where

$$N_k(\alpha,\beta,\gamma;t)=\int_{\sqrt{t}}^{t}(t-s)^{\alpha-1}s^{-\beta}L_{\gamma,k}(s)\,ds \tag{3.8}$$

and

$$E_K(t)=\frac{1}{\Gamma(\alpha)}\int_{\sqrt{t}}^{t}(t-s)^{\alpha-1}s^{-\beta}\varepsilon_K(s)\,ds. \tag{3.9}$$

In view of (3.5), we have

$$E_K(t)=O(N_K(\alpha,\beta,\gamma;t)) \qquad \text{as } t\to\infty \tag{3.10}$$

Further progress now depends on the asymptotic expansion of the integral $N_k(\alpha,\beta,\gamma;t)$.

Since $N_0(\alpha,\beta,\gamma;t)=L_\alpha(\beta,\gamma;t)$, Lemma 3 gives

$$N_0(\alpha,\beta,\gamma;t)\sim t^{\alpha-\beta}\sum_{r=o}^{\infty}\binom{\gamma}{r}\Omega_r(\alpha,\beta)(\ln t)^{\gamma-r} \qquad \text{as } t\to\infty. \tag{3.11}$$

In (3.8) we make the substitution $s=ut$. Then

$$N_k(\alpha,\beta,\gamma;t)=t^{\alpha-\beta}\int_{1/\sqrt{t}}^{1}(1-u)^{\alpha-1}u^{-\beta}L_{\gamma,k}(tu)\,du. \tag{3.12}$$

On the path of integration $t^{-1/2}\leqq u\leqq 1$, we have $|(\ln u)/\ln t|\leqq\frac{1}{2}$. Hence, for every integer $K\geqq 1$,

$$\ln\left(1+\frac{\ln u}{\ln t}\right)=\sum_{k=1}^{K}\frac{(-1)^k}{k}\frac{(\ln u)^k}{(\ln t)^t}+O\left(\frac{(\ln u)^{K+1}}{(\ln t)^{K+1}}\right). \tag{3.13}$$

This together with (2.11) gives

$$L_{\gamma,k}(tu)=L_{\gamma,k}(t)+(\gamma-k)L_{\gamma-1,k}(t)(\ln u)+O(L_{\gamma-2,k-1}(t)(\ln u)), \tag{3.14}$$

for $k=1,2,\cdots$. The constant involved in the O-symbol is independent of t and u.

We now substitute (3.14) in (3.12), and obtain from (2.7) and Lemma 2

$$N_k(\alpha, \beta, \gamma; t) = t^{\alpha-\beta}[\Omega_0(\alpha, \beta)L_{\gamma,k}(t)+(\gamma-k)\Omega_1(\alpha, \beta)L_{\gamma-1,k}(t)+O(L_{\gamma-2,k-1}(t))], \tag{3.15}$$

for $k=1, 2, \cdots$, as $t\to\infty$. The O-terms from (2.5) are included in that appearing in (3.15). On account of (3.11) and (3.15), we have

$$E_K(t)=O(t^{\alpha-\beta}L_{\gamma,K}(t)) \quad \text{as } t\to\infty. \tag{3.16}$$

Combination of (3.6), (3.7) and (3.16) yields

$$I^\alpha\phi(t)=\frac{1}{\Gamma(\alpha)}\sum_{k=0}^{K-1} c_kN_k(\alpha, \beta, \gamma; t)+O(t^{\alpha-\beta}L_{\gamma,K}(t)). \tag{3.17}$$

Substituting (3.11) and (3.15) in (3.17) and rearranging the terms, we obtain the desired result (3.2).

THEOREM 4. *Let $\phi(t)$ be locally integrable on* $[0, \infty)$ *and satisfy* (3.1) *with* $\beta=1$. *Then as* $t\to\infty$,

$$t^{1-\alpha}I^\alpha\phi(t)\sim\frac{1}{\Gamma(\alpha)}\int_0^t \phi(s)\,ds + \frac{(\ln t)^\gamma}{\Gamma(\alpha)}\left[d_0+d_1\frac{\ln\ln t}{\ln t}+d_2\frac{1}{(\ln t)}+d_3\frac{(\ln\ln t)^2}{(\ln t)^2}+d_4\frac{\ln\ln t}{(\ln t)^2}+\cdots\right], \tag{3.18}$$

where $d_0=c_0\Lambda_0(\alpha)$, $d_1=c_1\Lambda_0(\alpha)$, $d_2=c_0\Lambda_1(\alpha)$, $d_3=c_2\Lambda_0(\alpha)$ *and* $d_4=c_1(\gamma-1)\Lambda_1(\alpha)$; $\Lambda_0(\alpha)$ *and* $\Lambda_1(\alpha)$ *being the constants given in* (2.21) (*with* $k=0, 1$).

The proof is similar to that used in Theorems 2 and 3.

4. Application to the integral equation (1.4). The results obtained in the preceding section will now be applied to the solution of the integral equation (1.4). In terms of the fractional integral operator I^α, equation (1.4) may be written as

$$\phi(t)=I^{1/2}[f(t)-\phi^n(t)], \quad n\geqq 1, \tag{4.1}$$

where $f(t)$ satisfies (1.5), i.e.,

$$f(t)\sim\gamma_0t^{-a_0}+\gamma_1t^{-a_1}+\cdots, \quad \text{as } t\to\infty. \tag{4.2}$$

We recall that for $\alpha>0$, $\beta>0$,

$$I^\alpha I^\beta f=I^{\alpha+\beta}f. \tag{4.3}$$

Thus, applying $I^{1/2}$ on both sides of (4.1) gives

$$\int_0^t [f(s)-\phi^n(s)]\,ds=I^{1/2}\phi(t), \quad n\geqq 1. \tag{4.4}$$

We assume, as did Olmstead and Handelsman, that $\phi(t)$ has an asymptotic behavior of the form

$$\phi(t)\sim c_0t^{-\beta}(\ln t)^\gamma+\cdots \quad \text{as } t\to\infty. \tag{4.5}$$

The large-t behavior of $I^{1/2}\phi(t)$ is given in § 2. The behavior of the integral on the left-hand side of (4.4) can be obtained by simple computation. Upon balancing the terms of (4.4) asymptotically, we find $\beta=\frac{1}{2}$, $\gamma=-1$ and $c_0=\sqrt{\pi}$ when $n=2$ and $a_0>1$. This is precisely the result (1.6) given in [5, eq. (2.10)].

It is interesting to note that when $n=2$ and $a_0>1$, we have $I^{1/2}\phi(t)\to 0$ as $t\to 0$. Hence it follows from (4.4) that

$$\varepsilon_0=\int_0^\infty [f(s)-\phi^n(s)]\,ds=0. \tag{4.6}$$

One can in fact show that (4.6) holds for $1\leqq n\leqq 2$ and $a_0>1$. This property of the solution marks the distinct difference between cases in which $1\leqq n\leqq 2$ and those in which $n>2$, since in the latter cases it was shown by Handelsman and Olmstead [4, p. 381] that $\varepsilon_0>0$. In view of (4.6), we may write (4.4) as

$$\int_t^\infty [\phi^n(s)-f(s)]\,ds=I^{1/2}\phi(t), \qquad 1\leqq n\leqq 2. \tag{4.7}$$

Returning to the case $n=2$ and $a_0>1$, we have

$$\phi(t)\sim\sqrt{\pi}\,t^{-1/2}/\ln t \qquad \text{as } t\to\infty, \tag{4.8}$$

from which one would expect that $\phi(t)$ may have an asymptotic expansion of the form (1.3). However any attempt to satisfy (4.7), as $t\to\infty$, with the second term being of the form $c_1t^{-1/2}(\ln t)^\sigma$ leads to a contradiction in equating the coefficient c_1 (although the exponents seem to agree on both sides). By assuming $\phi(t)$ to have an asymptotic behavior of the form

$$\phi(t)\sim\frac{\sqrt{\pi}}{\sqrt{t}\ln t}\left[1+c_1\frac{\ln\ln t}{\ln t}+c_2\frac{(\ln\ln t)^2}{(\ln t)^2}+\cdots\right], \tag{4.9}$$

we obtain $c_1=2[\psi(1)-\psi(\frac{1}{2})]$ and $c_2=c_1^2/\pi$.

The case $n=\frac{3}{2}$ and $a_0\geqq\frac{3}{2}$ also deserves an attention. The leading term of the asymptotic solution $\phi(t)$ is given by

$$\phi(t)\sim\frac{1}{36\pi}t^{-1}(\ln t)^2+\cdots \qquad \text{as } t\to\infty, \tag{4.10}$$

see [5, eq. (2.6)]. The succeeding terms again involve the function $\ln\ln t$. We obtain

$$\phi(t)\sim\frac{1}{36\pi}\frac{(\ln t)^2}{t}\left[1+c_1\frac{\ln\ln t}{\ln t}+c_2\frac{(\ln\ln t)^2}{(\ln t)^2}+\cdots\right], \tag{4.11}$$

where $c_1=4(\psi(1)-\psi(\frac{1}{2})-2)$ and $c_2=c_1^2/4$.

REFERENCES

[1] NEIL BERGER AND RICHARD HANDELSMAN, *Asymptotic evaluation of fractional integral operators with applications*, this Journal, 6 (1975), pp. 766–773.

[2] NORMAN BLEISTEIN, *Asymptotic expansions of integral transforms of functions with logarithmic singularities*, this Journal, 8 (1977), pp. 655–672.

[3] J. J. DORNING, B. NICOLAENKO AND J. K. THURBER, *An integral identity due to Ramanujan which occurs in neutron transport theory*, J. Math. Mech., 19 (1969), pp. 429–449.

[4] RICHARD A. HANDELSMAN AND W. E. OLMSTEAD, *Asymptotic solution to a class of nonlinear Volterra integral equations*, SIAM J. Appl. Math., 22 (1972), pp. 373–384.

[5] W. E. OLMSTEAD AND RICHARD A. HANDELSMAN, *Asymptotic solution to a class of nonlinear Volterra integral equations II*, Ibid., 30 (1976), pp. 180–189.

[6] K. STEWARTSON, *On asymptotic expansions in the theory of boundary layers*, J. Math. and Phys., 36 (1957), pp. 172–191.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0008 $01.00/0

ZEROS OF THE SOLUTIONS OF FIRST ORDER FUNCTIONAL DIFFERENTIAL EQUATIONS*

CLEMENT McCALLA†

Abstract. It is shown that all solutions of a class of first order autonomous functional differential equations are oscillatory if the fundamental solution has a first zero T_1. If the fundamental solution does not have any zeros, we show that every solution with positive initial data is nonoscillatory. Upper and lower bounds are given for the distance between successive zeros in terms of T_1 and the delay a.

1. Introduction. A number of authors have studied the oscillatory behavior of the solutions of first order functional differential equations, for example [1], [2], [6], [8], [9], [10] and [11]. In particular, Myshkis [8, Chap. 4] considers the non-autonomous first order functional differential equation of "stable type"

$$x'(t) = -\int_0^{\Delta(t)} x(t-s)\,d_sK(t,s), \qquad t > t_0, \tag{1.1}$$

with continuous initial function $\phi(t)$, continuous delay function $\Delta(t) \geqq 0$, kernel $K(t,s)$ for fixed t is of bounded variation in s and does not decrease, $K(t,0)=0$, $K(t,s)=K(t,\infty)$ for $s > \Delta(t)$ and $K(t,s)$ satisfies the condition of mean continuity. Defining

$$M_0 = \sup_{t\in[t_0,\infty)} \bigvee_{s=0}^{\Delta(t)} K(t,s) \quad \text{and} \quad \Delta_0 = \sup_{t\in[t_0,\infty)} \Delta(t),$$

and assuming that $\phi(t) \geqq 0$, $0 < M_0 < \infty$ and $\Delta_0 M_0 < 1$, Myshkis shows that every semicycle of the solution is large and with length greater than M_0^{-1}. In addition, Myshkis [9] considers the equation

$$x'(t) = -M(t)x(t-\Delta(t))$$

where $M(t) \geqq 0$ is continuous. Defining

$$m_0 = \inf_{t\in[t_0,\infty)} M(t) \quad \text{and} \quad \tau_0 = \inf_{t\in[t_0,\infty)} \Delta(t)$$

and assuming that $\Delta_0 < \infty$ and $\tau_0 m_0 > e^{-1}$, Myshkis shows that every solution is oscillatory.

More recently, a great deal of attention has been centered on the foundations of the theory of functional differential equation; see for example Hale [5]. In this paper we use a representation of solutions (2.2) and an addition type formula for the solution (2.3) to obtain results on the zeros of a class of autonomous first order functional differential equations assuming only the existence or nonexistence of a first zero of the fundamental solution. If the fundamental solution has a first zero "T_1," we show that all solutions are oscillatory and obtain upper and lower bounds on the distance between successive zeros in terms of "T_1" and the delay "a." On the other hand, if the fundamental solution has no zeros, we show that every solution with positive initial data can have at most one zero and that the distance between successive zeros of any oscillatory solution is less than the delay "a." Finally, we exhibit sufficient conditions for the existence and nonexistence of a first zero of the

* Received by the editors March 18, 1976, and in final revised form February 14, 1977.

† Department of Mathematics, Howard University, Washington, DC 20059. This work was supported in part by the National Science Foundation under Grant SER75-09043.

fundamental solution of some differential equations in terms of the coefficients appearing in the equations and the delay "a."

2. Preliminaries. We consider the equation

$$(2.1)\qquad \begin{aligned} x'(t) &= \sum_{i=0}^{N} A_i x(t+\theta_i) + \int_0^a A(\theta)x(t-\theta)\,d\theta, \qquad t>0, \\ x(t) &= h(t) \qquad -a \leqq t \leqq 0, \end{aligned}$$

where A_i, θ_i are constants, $a>0, -a=\theta_N<\cdots<\theta_1<\theta_0=0$, $h(0)$ is given, $h(\cdot)\in L^p(-a,0)$ for $p\geqq 1$, and $A(\cdot)\in L^q(0,a)$ where $q=p/(p-1)$. In addition we assume that one of the following conditions is satisfied:

(C_1): $A_i\leqq 0$, $i=1,\cdots,N-1$, $A_N<0$, and $A(\theta)\leqq 0$ a.e. with respect to Lebesgue measure on $[0,a]$.

(C_2): $A_i\leqq 0$, $i=1,\cdots,N$, $A(\theta)\leqq 0$ a.e. with respect to Lebesgue measure on $[0,a]$, and $A(\theta)$ is not a null function on any interval $[a-\delta, a]$ for $0<\delta\leqq\delta_0$ and some $\delta_0>0$.

Theorems giving the existence, uniqueness and the representation of the solution to (2.1) can be found in [3] and [5]. The solution $x(t)$ will be (absolutely) continuous for $t>0$, continuous on the right at $t=0$, and can be given in the form

$$(2.2)\qquad \begin{aligned} x(t) = \Phi(t)h(0) &+ \sum_{i=1}^{N} A_i \int_{\theta_i}^{0} \Phi(t+\theta_i-\alpha)h(\alpha)\,d\alpha \\ &+ \int_{-a}^{0} d\alpha \int_{-\alpha}^{a} d\theta\, A(\theta)\Phi(t-\theta-\alpha)h(\alpha) \end{aligned}$$

where $\Phi(t)$, the fundamental solution of (2.1), satisfies (2.1) with initial data $\Phi(0)=1$, $\Phi(t)=0$ for $-a\leqq t<0$. An expression for $\Phi(t)$ can be found in [7].

The following lemma follows from the representation of solutions (2.2) and the observation that the translate of a solution of the autonomous equation (2.1) is a solution.

LEMMA 1.

$$(2.3)\qquad \begin{aligned} x(t+s) = \Phi(t)x(s) &+ \sum_{i=1}^{N} A_i \int_{\theta_i}^{0} \Phi(t+\theta_i-\alpha)x(s+\alpha)\,d\alpha \\ &+ \int_{-a}^{0} d\alpha \int_{-\alpha}^{a} d\theta\, A(\theta)\Phi(t-\theta-\alpha)x(s+\alpha) \end{aligned}$$

for $t, s\geqq 0$.

DEFINITIONS. (i) The solution $x(t)$ is said to be *oscillatory* if it changes sign on any interval $[\tau,\infty)$.

(ii) The interval $[\tau_1,\tau_2]$ is said to be a *semi-cycle* of the solution if $x(t)\neq 0$ on (τ_1,τ_2) and $x(\tau_1)=x(\tau_2)=0$.

(iii) The semi-cycle $[\tau_1,\tau_2]$ is said to be *large* if $\tau_2-\tau_1>a$; otherwise it is said to be *small.*

(iv) The solution $x(t)$ is said to have *first zero* t_1 if $t_1>0$, $x(t)\neq 0$ on $(0,t_1)$ and $x(t_1)=0$.

3. Main results.

THEOREM 1. *If the fundamental solution* $\Phi(t)$ *of* (2.1) *has a first zero* $T_1>0$, *then all nontrivial solutions are oscillatory and have semi-cycles of length less than* T_1+a.

Moreover if $T_1 > a$, all solutions with nontrival initial data of constant sign (including zero) have large semi-cycles, change sign at each zero t_n, and $T_1 < t_{n+1} - t_n < T_1 + a$.

Proof. Suppose that $x(t)$ is a nontrivial solution of (2.1) and that $x(t) > 0$ for $\tau < t < \tau + T_1 + a$ and $x(\tau + T_1 + a) \geqq 0$, where $\tau > 0$. From equation (2.4) we have that

$$\begin{aligned} x(T_1+\tau+a) = &-\sum_{i=1}^{N} |A_i| \int_{\theta_i}^{0} \Theta(T_1+\theta_i-\alpha)x(\tau+a+\alpha)\,d\alpha \\ &-\int_{-a}^{0} d\alpha \int_{-\alpha}^{a} d\theta |A(\theta)| \Phi(T_1-\theta-\alpha)x(\tau+a+\alpha) \\ &< 0. \end{aligned} \tag{3.1}$$

This is a contradiction and hence every nontrivial solution must have a zero on an interval of length $T_1 + a$.

Let $x(t)$ be a solution with initial data $h(\alpha) \geqq 0$ and $h(0) > 0$ (the proof is similar in the case $h(0) = 0$). We define $f(s) = x(t_1 + s)$ where the first zero $t_1 > 0$ exists by the first part of the theorem. From (2.3) we have that

$$\begin{aligned} f_1(s) = &-\sum_{i=1}^{N} |A_i| \int_{\theta_i}^{0} \Phi(s+\theta_i-\alpha)x(t_1+\alpha)\,d\alpha \\ &-\int_{-\alpha}^{0} d\alpha \int_{\alpha}^{a} d\theta |A(\theta)| \Phi(s-\theta-\alpha)x_1(t+\alpha), \end{aligned} \tag{3.2}$$

and thus $f_1(s) < 0$ for $0 < s < T_1$. Hence $f_1(s)$ has a zero in $[T_1, T_1 + a]$, that is, $x(t)$ has a second zero t_2 in $[t_1 + T_1, t_1 + T_1 + a]$. Moreover, $x(t)$ changes sign at t_2, since $T_1 > a$ implies that $\dot{x}(t_2) > 0$. At the nth step, we define $f_n(s) = x(t_n + s)$ and repeat the procedure to show that $x(t)$ has a zero t_{n+1} in $[t_n + T_1, t_n + T_1 + a]$.

Remarks 1. Professors R. D. Driver and J. A. Yorke have communicated to the author a proof of the first part of Theorem 1 independent of a representation of solutions and based upon a consideration of the function $F(t) = \log(\Phi(t)/x(t))$, $0 < t < T_1$ and the hypothesis that $x(t) > 0$ for $-a < t < T_1$.

2. The last part of Theorem 1 holds if we assume that the initial data $h(\alpha) \geqq 0$ ($\leqq 0$) a.e. with respect to Lebesgue measure on $[-a, 0]$ and that $h(\alpha)$ is not a null function. If $h(0) > 0$ (< 0), $0 < t_1 < T_1$ and if $h(0) \leqq 0$ ($\geqq 0$), $T_1 < t_1 < T_1 + a$.

THEOREM 2. *If the fundamental solution $\Phi(t)$ of* (2.1) *is positive on R^+, then solutions with nontrivial initial data of constant sign (including zero) have at most one zero on R^+ and hence are nonoscillatory. In addition, oscillatory solutions have small semi-cycles.*

Proof. Suppose that $x(t)$ is a solution with initial data $h(\alpha) \geqq 0$. If $h(0) = 0$, (2.2) implies that $x(t) < 0$ on $(0, \infty)$. If $h(0) > 0$ and $x(t)$ has no zeros, we are done; if $t_1 > 0$ is the first zero of $x(t)$, then (3.2) implies that $x(t_1 + s) < 0$ on (t_1, ∞). Consequently, if $y(t)$ is an oscillatory solution with consecutive zeros τ_1 and τ_2, then $\tau_2 - \tau_1 < a$, else the solution $z(s) = y(\tau_2 + s)$ would have initial data of constant sign and thus be nonoscillatory.

Example 1.

$$x'(t) = A_0 x(t) - |A_1| x(t-a), \qquad a|A_1| \geqq e^{A_0 a}. \tag{3.3}$$

The fundamental solution is given by

$$\Phi(t) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} e^{A_0(t-na)} |A_1|^n (t-na)_+^n, \tag{3.4}$$

where $(t)_+ = \max(t, 0)$. It is easy to see that $T_1 = a + |A_1|^{-1} e^{A_0 a} > a$. From a result of Myshkis [8], it follows that T_1 exists for $a|A_1| > e^{-1} e^{A_0 a}$ and in this case clearly $T_1 > a$.

Example 2.

$$x'(t) = -\int_0^a k^2 x(t-\theta)\,d\theta, \qquad k > 0, \tag{3.5}$$

where $ka < \pi/2$ and $\cos(2ka) + \frac{1}{2}ka \sin(ka) < 0$, i.e.,

$$1.004081\cdots < ka < 1.570796\cdots. \tag{3.6}$$

The fundamental solution is given by

$$\Phi(t) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} k^{2n} D_x^n f(x, t-na)\Big|_{x=0} H(t-na), \tag{3.7}$$

where

$$f(x, t) = \cos(k^2 + x)^{1/2} t$$

and $H(t)$ is the unit Heaviside step function. From (3.6) it follows that $\Phi(t) > 0$ for $0 \leqq t \leqq a$ and that $\Phi(2a) < 0$. Hence T_1 exists and $T_1 > a$. Note that $T_1 = \pi/(2k) \leqq a$ if $ka \geqq \pi/2$.

So far we have considered examples in which the fundamental solution has a first zero T_1. We will now consider examples in which the fundamental solution $\Phi(t)$ is positive on R^+. If the characteristic function $\Delta(\lambda)$ corresponding to (2.1) has a real root α, then (2.1) has a nonoscillatory exponential solution $x(t) = e^{\alpha t}$. Hence $\Phi(t) > 0$ for all $t \geqq 0$. For if $\Phi(t)$ has a first zero, then from Theorem 1 every nontrivial solution would be oscillatory and this is a contradiction.

Example 3.

$$x'(t) = -\int_0^a k^2 x(t-\theta)\,d\theta, \qquad k > 0, \tag{3.8}$$

where $(ka)^2 < \mu^2(e^\mu - 1)^{-1} = 0.67102\cdots$ and $\mu = 1.593624\cdots$ is the value of $x > 0$ maximizing the function $f(x) = x^2(e^x - 1)^{-1}$. The characteristic equation

$$\Delta(\lambda) = \lambda + k^2 \int_0^a e^{-\lambda\theta}\,d\theta = 0 \tag{3.9}$$

has a real root $\alpha \in (-\mu a^{-1}, 0)$, since $\Delta(-\mu a^{-1}) < 0$ and $\Delta(0) > 0$. Hence the fundamental solution $\Phi(t)$ given by (3.7) is positive on R^+.

Example 4.

$$x'(t) = -\sum_{i=0}^{N} |A_i| x(t+\theta_i) - \int_0^a |A(\theta)| x(t-\theta)\,d\theta, \qquad Kae < 1, \tag{3.10}$$

where

$$K = \sum_{i=0}^{N} |A_i| + \int_0^a |A(\theta)|\,d\theta > 0.$$

Driver [4] has obtained an asymptotic characterization of the solution of the "differential equation with small delay" (3.10) from which one can conclude that the solutions of (3.10) practically never oscillate. The interested reader should consult [4] for further details. The characteristic equation

$$\Delta(\lambda)=\lambda+\sum_{i=0}^{N}|A_i|\,e^{\lambda\theta_i}+\int_0^a |A(\theta)|\,e^{-\lambda\theta}\,d\theta=0 \tag{3.11}$$

has a real root $\alpha\in(-a^{-1},0)$, since $\Delta(-a^{-1})<0$ and $\Delta(0)>0$. Hence the fundamental solution $\Phi(t)$ of (3.10) is positive on R^+.

Acknowledgment. The author wishes to thank the reviewers for their very helpful remarks.

REFERENCES

[1] G. BIRKHOFF AND K. KOTIN, *Asymptotic behavior of first order linear differential-delay equations*, J. Math. Anal. Appl., 13 (1966), pp. 8–18.

[2] J. BUCHANAN, *Bounds on the growth of a class of oscillatory solutions of* $y'(x)=my(x-d(x))$ *with bounded delays*, SIAM J. Appl. Math., 27 (1974), pp. 18–28.

[3] M. C. DELFOUR AND S. K. MITTER, *Hereditary differential systems with constant delays*, J. Differential Equations, 12 (1972), pp. 213–235; 18 (1975), pp. 18–28.

[4] R. D. DRIVER, *Linear differential systems with small delays*, Ibid., 21 (1976), pp. 148–166.

[5] J. K. HALE, *Functional Differential Equations*, Springer-Verlag, New York, 1971.

[6] J. C. LILLO, *Oscillatory solutions of* $y'(x)=m(x)y(x-n(x))$, J. Differential Equations, 6 (1969), pp. 1–35.

[7] C. MCCALLA, *The exact solutions of some functional differential equations*, Dynamical Systems: An International Symposium, Vol. 2, L. Cesari, J. K. Hale and J. P. La Salle, eds., Academic Press, New York, 1975, pp. 163–168.

[8] A. D. MYSHKIS, *Linear Differential Equations with Retarded Argument*, Gittl, Moscow, 1951; German transl. Veb. Deutscher Verlag, Berlin, 1955.

[9] ———, *On the solutions of linear homogeneous differential equations of the first order and stable type with retarded arguments*, Mat. Sb., 28 (1951), pp. 641–658. (In Russian.)

[10] E. WINSTON, *Comparison theorems for scalar delay differential equations*, J. Math. Anal. Appl., 29 (1970), pp. 455–563.

[11] E. M. WRIGHT, *A non-linear difference-differential equation*, J. Reine Angew. Math., 194 (1955), pp. 66–87.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0009 $01.00/0

S-ORTHOGONAL PROJECTION OPERATORS AS ASYMPTOTIC SOLUTIONS OF A CLASS OF MATRIX DIFFERENTIAL EQUATIONS*

ERKKI OJA†

Abstract. A class of nonlinear autonomous matrix differential equations is considered. Two special cases of this class yield the differential equations of adaptive network models which were initially introduced in connection with idealized neuron networks. It is shown that these equations exhibit an asymptotic behavior that is intimately related with the matrix form of the Gram–Schmidt orthogonalization algorithm in a real vector space, with respect to an inner product with an arbitrary positive definite symmetric weight matrix S.

1. Introduction. Several applications of pattern recognition, associative memories, and theory of learning systems are concerned with the problem of constructing projection operators on specified subspaces. Such operators frequently result as asymptotic states or goals of learning in certain adaptive processes (for a review, see [3]). Recently, Kohonen [3], [5] has suggested a dynamical network model whose overall transfer matrix has interesting projection properties. The transfer matrix is governed by the equation

$$\frac{d\phi}{dt} = -\alpha\phi^2 aa^T\phi^2, \qquad t \geqq 0,$$

where ϕ is an $n \times n$ matrix function of t, $a \in R^n$ is a vector and $\alpha > 0$ is a scalar. In [5] it is shown for the corresponding difference equation that the solution, starting from a projection matrix, tends asymptotically to a projection matrix on a given subspace, when the input of the network, or vector a, is suitably chosen.

In this paper, the above mathematical result is applied to matrix differential equations of a more general type, to appear as equations (12) and (25) in following sections. No physical implementations are here regarded; however, equations of this class may represent the effect of certain variations in the basic network model introduced in [3]. It will be shown that the relation between the initial and asymptotic solutions is intimately connected with a step in the matrix form of the Gram–Schmidt orthogonalization algorithm. Thus it will be seen that, in a sense, this class of equations is a continuous counterpart of the Gram–Schmidt algorithm in R^n with an inner product having a weight matrix that is either the unit matrix or a positive definite arbitrary matrix.

2. Gram–Schmidt construction of an S-orthogonal basis by means of nonorthogonal projection matrices. This section is a brief presentation of well-known properties of Gram–Schmidt orthogonalization, put in a form that allows comparison with the asymptotic properties of matrix differential equations under study.

Let $\{a_1, a_2, \cdots\}$ be an arbitrary sequence of vectors in R^n. Consider the unnormalized Gram–Schmidt algorithm

$$(1) \qquad \begin{aligned} \tilde{a}_1 &= a_1, \\ \tilde{a}_k &= a_k - \sum_{j \in S_k} \frac{\tilde{a}_j^T a_k}{\tilde{a}_j^T \tilde{a}_j} \tilde{a}_j, \qquad k > 1 \end{aligned}$$

* Received by the editors September 9, 1976.

† Department of Technical Physics, Helsinki University of Technology, SF-02150 Espoo 15, Finland.

where

(2) $$S_k = \{j | 1 \leqq j \leqq k-1 \text{ and } \tilde{a}_j \neq 0\},$$

yielding for every k the orthogonal basis $\{\tilde{a}_1, \cdots, \tilde{a}_k\}$ of the subspace $L(a_1, \cdots, a_k)$ spanned by the vectors $a_1, \cdots, a_k$. It is shown e.g. in [1, p. 37] that formulae (1), (2) may be expressed in an equivalent matrix form as follows:

(3) $$\tilde{a}_k = \phi_{k-1} a_k,$$

(4) $$\phi_k = \begin{cases} \phi_{k-1} - \dfrac{\phi_{k-1} a_k a_k^T \phi_{k-1}}{a_k^T \phi_{k-1} a_k} & \text{if } \phi_{k-1} a_k \neq 0, \\ \phi_{k-1} & \text{otherwise,} \end{cases}$$

$$\phi_0 = I.$$

Now ϕ_k is the orthogonal projection matrix on the orthogonal complement of $L(a_1, \cdots, a_k)$; consequently, every ϕ_k is symmetric and idempotent.

Let now R^n be equipped with an inner product defined by

(5) $$(x, y)_S = x^T S y = y^T S x,$$

where S is a symmetric positive definite matrix. Then a sequence of vectors $\{h_1, h_2, \cdots\}$ satisfying

(6) $$(h_i, h_j)_S = 0 \qquad \text{if } i \neq j$$

will be generated recursively from the vectors $\{a_1, a_2, \cdots\}$ by the following algorithm:

(7) $$h_1 = a_1,$$

(8) $$h_k = a_k - \sum_{j \in S'_k} \frac{(h_j, a_k)_S}{(h_j, h_j)_S} h_j = a_k - \sum_{j \in S'_k} \frac{h_j^T S a_k}{h_j^T S h_j} h_j = \Big(I - \sum_{j \in S'_k} \frac{h_j h_j^T S}{h_j^T S h_j}\Big) a_k$$

where

(9) $$S'_k = \{j | 1 \leqq j \leqq k-1 \text{ and } h_j \neq 0\}.$$

Vectors $\{h_1, h_2, \cdots\}$ are S-orthogonal, or conjugate with respect to S; thus the above algorithm may be termed an S-orthogonalizing algorithm. If the matrix mapping a_k into h_k is denoted by ψ_{k-1}, it is evident from (8) that the subsequent matrices ψ_k and ψ_{k-1} are related as follows:

(10) $$\psi_k = \begin{cases} \psi_{k-1} - \dfrac{h_k h_k^T S}{h_k^T S h_k} & \text{if } h_k \neq 0, \\ \psi_{k-1} & \text{otherwise,} \end{cases}$$

with

$$\psi_0 = I.$$

Substituting $h_k = \psi_{k-1} a_k$ in (10) yields further

(11) $$\psi_k = \psi_{k-1} - \frac{\psi_{k-1} a_k a_k^T \psi_{k-1}^T S}{a_k^T \psi_{k-1}^T S \psi_{k-1} a_k} \quad \text{if } h_k \neq 0.$$

Because of the well-known properties of Gram–Schmidt orthogonalization, the matrices ψ_k map each vector a_{k+1} into a subspace S-orthogonal to $L(a_1, \cdots, a_k)$ or orthogonal to $L(Sa_1, \cdots, Sa_k)$ in the Euclidean sense.

The two recursions (3)–(4) and (10)–(11) will be related to the behavior of matrix differential equations in the following section.

3. Matrix differential equations and Gram–Schmidt orthogonalization. In the sequel, let $\phi(\cdot)$ denote a continuous and differentiable matrix function defined on the set of nonnegative reals R^+. Let $a \in R^n$ be a real vector and $P(\cdot)$ and $Q(\cdot)$ be functions expressible as power series with real scalar coefficients. Consider the following differential equation:

$$\frac{d\phi}{dt} = -P(\phi)aa^TQ(\phi^T), \qquad t \in R^+. \tag{12}$$

Throughout this paper the assumption is made that (12) is autonomous, i.e., neither the vector a nor the coefficients of $P(\cdot)$ and $Q(\cdot)$ are dependent on t. The case in which this assumption is not valid is treated elsewhere [6].

A special case of (12) has earlier been discussed by Kohonen [3] and by Kohonen and Oja [5]:

$$\frac{d\phi}{dt} = -\alpha\phi^2 aa^T\phi^2, \qquad \alpha > 0, \tag{13}$$

with initial condition

$$\phi(0)^2 = \phi(0) = \phi(0)^T. \tag{14}$$

The solution is

$$\phi(t) = \phi(0) + \phi(0)aa^T\phi(0)\varphi(t) \tag{15}$$

with

$$\varphi(t) = \{[3\alpha(a^T\phi(0)a)t + 1]^{-1/3} - 1\}(a^T\phi(0)a)^{-1} \tag{16}$$

if

$$\phi(0)a \neq 0,$$

i.e., $\varphi(t)$ is a scalar valued function of t with initial value $\varphi(0) = 0$ and asymptotic value $\lim_{t\to\infty} \varphi(t) = -(a^T\varphi(0)a)^{-1}$. Then the solution of (13)–(14) converges to

$$\lim_{t\to\infty} \phi(t) = \phi(0) - \phi(0)aa^T\phi(0)(a^T\phi(0)a)^{-1}. \tag{17}$$

On the other hand, if $\phi(0)a = 0$, then $\varphi(t)$ is identically zero, and $\phi(t)$ remains equal to $\phi(0)$.

The above solution can now be generalized to the present case of (12).

THEOREM 1. *Let $\phi(0)$ be an idempotent symmetric $n \times n$ matrix and $a \in R^n$ a constant vector. Let $P(\cdot)$ and $Q(\cdot)$ satisfy the following*:

1. *As functions of a scalar variable, they possess convergent power series expansions on the interval* $[0, 1]$, *vanishing at zero*;
2. *on the interval* $(0, 1]$, *the function* $P(\cdot)Q(\cdot)$ *is positive.*

Then the solution of (12) *tends asymptotically to the matrix* (17), *if* $\phi(0)a \neq 0$; *if* $\phi(0)a = 0$, *then* $\phi(t)$ *is identically equal to* $\phi(0)$.

Proof. It is first shown that the solution of (12) is of the form (15). Substitution yields

$$\phi(t)a=\phi(0)a[1+\varphi(t)a^T\phi(0)a]=\phi(0)a\sigma(t)$$

where

$$\sigma(t)=1+\varphi(t)a^T\phi(0)a. \tag{18}$$

Furthermore,

$$\begin{aligned}\phi(t)^2a&=\phi(t)\phi(0)a\sigma(t)\\&=\phi(0)^2a\sigma(t)+\varphi(t)a^T\phi(0)a\phi(0)a\sigma(t)\\&=\phi(0)a\sigma(t)^2\end{aligned}$$

and by induction

$$\phi(t)^ma=\phi(0)a\sigma(t)^m,\qquad m\geqq 1.$$

$P(\cdot)$ may be represented in the form

$$P(\cdot)=\sum_{j=0}^{\infty}\pi_j(\cdot)^j \tag{19}$$

with π_j real for each j and $\pi_0=0$ because $P(\cdot)$ vanishes at zero. This yields

$$P(\phi)a=\sum_{j=1}^{\infty}\pi_j\phi^ja=\phi(0)a\sum_{j=1}^{\infty}\pi_j\sigma^j=\phi(0)aP(\sigma) \tag{20}$$

on an interval J of t where $0\leqq\sigma(t)\leqq 1$ and $0\leqq\rho[\phi(t)]\leqq 1$, $\rho[\cdot]$ denoting the spectral radius of a square matrix. It is evident that the point $t=0$ is included in J since (15) and (18) show that $\sigma(0)=1$, and the spectral radius of an idempotent matrix is at most 1.

Forming $Q(\phi)a$ in a similar fashion yields the following expression:

$$P(\phi)aa^TQ(\phi^T)=\phi(0)aa^T\phi(0)P(\sigma)Q(\sigma), \tag{21}$$

and substitution into (12) yields the scalar differential equation

$$\frac{d\sigma}{dt}=-(a^T\phi(0)a)P(\sigma)Q(\sigma),\qquad t\in J, \tag{22}$$

with initial value

$$\sigma(0)=1. \tag{23}$$

If $\phi(0)a=0$, the solution is trivially $\sigma(t)=1$ everywhere, implying that $\varphi(t)$ is identically zero and $\phi(t)$ is equal to $\phi(0)$. Therefore, consider the nontrivial case $\phi(0)a\neq 0$. The assumptions made on $P(\cdot)$ and $Q(\cdot)$ imply now that $P(\sigma)Q(\sigma)$ is continuous and satisfies a Lipschitz condition with respect to σ on $[0, 1]$. Therefore, it follows from the Picard–Lindelöf theorem [2] that on J equation (22) has a unique solution corresponding to the initial value (23). Denote this unique solution by $\sigma^*(t)$. Since the positiveness of $P(\cdot)Q(\cdot)$ implies that the right side of (22) is strictly negative as long as $\sigma\in(0, 1]$, it follows that on J the solution $\sigma^*(t)$ is monotonically nonincreasing. On the other hand, since $P(\cdot)Q(\cdot)$ vanishes at zero, equation (22) has the trivial solution $\sigma(t)=0$ on R^+. Uniqueness implies now that $\sigma^*(t)$ cannot vanish and therefore it can be concluded that J is arbitrarily long. Since $\sigma^*(t)$ is now nonincreasing and bounded from below on R^+, it tends to some number $\bar{\sigma} \in [0, 1]$; however, $\bar{\sigma}$ must be a root of $P(\cdot)Q(\cdot)$, which implies that $\bar{\sigma} = 0$.But then, by (18), $\lim_{t\to\infty} \varphi(t) = -(a^T\phi(0)a)^{-1}$, and substitution into (15) establishes the theorem.

In many cases a complete solution of (12) on R^+, corresponding to a symmetric idempotent initial matrix, is naturally obtained by putting (22) in the form

$$\int_1^{\sigma} \frac{d\omega}{P(\omega)Q(\omega)} = -(a^T\phi(0)a)t, \tag{24}$$

by integrating and by solving for σ as a function of t.

It is, however, the asymptotic solution (17) that is especially interesting because of its relation with $\phi(0)$. Equation (17) should be compared with (4), yielding the general step in the Gram–Schmidt algorithm. As pointed out e.g. in [5], setting $\phi(0) = \phi_{k-1}$ and $a = a_k$ in (13) results in the asymptotic behavior $\lim_{t\to\infty} \phi(t) = \phi_k$, which corresponds to a step in algorithm (4). Theorem 1 above generalizes this result to equations of the type of (12).

It will now be shown that a generalization of (12) produces an asymptotic behavior that is intimately related with the S-orthogonalization procedure of (10). Instead of (12), consider the following matrix differential equation:

$$\frac{d\psi}{dt} = -P(\psi)aa^TQ(\psi^T)SR(\psi), \qquad t \in R^+ \tag{25}$$

where $\psi(\cdot)$ is continuous and differentiable on R^+, $a \in R^n$ is constant and $P(\cdot)$, $Q(\cdot)$, and $R(\cdot)$ are again expressible as power series with real scalar coefficients. S is a positive definite symmetric matrix independent of t.

It has been pointed out by Kohonen [4] that a special case of (25), with $P(\psi) = \psi^2$, $Q(\psi) = R(\psi) = \psi$, appears as the differential equation of an adaptive network model that is of interest in the mathematical treatment of idealized neuron networks.

To proceed with the solution of (25), consider the similarity transformation

$$\Gamma(t) = S^{1/2}\psi(t)S^{-1/2} \tag{26}$$

$$\psi(t) = S^{-1/2}\Gamma(t)S^{1/2}, \qquad t \in R^+. \tag{27}$$

Denote further

$$b = S^{1/2}a, \qquad a = S^{-1/2}b. \tag{28}$$

Substitution into (25) yields

$$\begin{aligned} \frac{d\Gamma}{dt} &= S^{1/2}\frac{d\psi}{dt}S^{-1/2} \\ &= -S^{1/2}P(\psi)S^{-1/2}bb^TS^{-1/2}Q(\psi^T)SR(\psi)S^{-1/2} \\ &= -P(\Gamma)bb^TQ(\Gamma^T)R(\Gamma) \end{aligned} \tag{29}$$

since naturally $P(\Gamma) = S^{1/2}P(\psi)S^{-1/2}$, etc.

The only basic difference between (29) and (12) is that on the right side of (29) there appear matrix products of the form $\Gamma^T\Gamma$, due to the factor $Q(\Gamma^T)R(\Gamma)$, instead of mere powers of Γ^T and Γ. If, however, Γ were symmetric, then this difference would vanish; for suitable initial conditions, Theorem 1 might then be used to construct the asymptotic solution of (29) and, consequently, that of (25). This is established in the following.

THEOREM 2. *In* (25), *let* $P(\cdot)$, $Q(\cdot)$, *and* $R(\cdot)$ *satisfy Condition* 1 *imposed in Theorem* 1, *and let* $P(\cdot)Q(\cdot)R(\cdot)$ *be positive on the interval* (0, 1]. *Let a be a constant vector and S a constant positive definite symmetric matrix. Furthermore, let* $\psi(0)$ *be idempotent and such that* $S\psi(0)$ *is symmetric. Then the asymptotic solution of* (25) *is given by*

$$\lim_{t\to\infty} \psi(t) = \psi(0) - \frac{\psi(0)aa^T\psi(0)^T S}{a^T\psi(0)^T S\psi(0)a}. \tag{30}$$

Proof. Equation (26) implies that $\Gamma(0)$ is symmetric if and only if $S\psi(0)$ is symmetric. Likewise, $\Gamma(0)$ is idempotent if and only if $\psi(0)$ is idempotent. Then the initial conditions of (29) are exactly those of Theorem 1; a construction similar to that in the proof of Theorem 1 guarantees that (29) now has a solution of the form

$$\Gamma(t) = \Gamma(0) + \gamma(t)\Gamma(0)bb^T\Gamma(0), \tag{31}$$

with $\gamma(t)$ a scalar-valued function. Then Theorem 1 is applicable even in the present case, implying that $\Gamma(t)$ tends to

$$\lim_{t\to\infty} \Gamma(t) = \Gamma(0) - \frac{\Gamma(0)bb^T\Gamma(0)}{b^T\Gamma(0)b}. \tag{32}$$

Employing the similarity transformation (26) yields

$$\lim_{t\to\infty} \psi(t) = \psi(0) - \frac{\psi(0)aa^T S\psi(0)}{a^T S\psi(0)a}, \tag{33}$$

which is equivalent to (30) when use is made of the facts that $S\psi(0)$ is symmetric and $\psi(0)$ is idempotent. This concludes the proof of Theorem 2.

Comparing (30) with the recursion formula (11), it is evident that setting $\psi(0) = \psi_{k-1}$ and $a = a_k$ yields $\lim_{t\to\infty} \psi(t) = \psi_k$ or the next step in the recursion (11). Thus (11) is in a sense the discrete counterpart of the differential equation (25), one discrete step there corresponding to integration over an infinite interval in the continuous case. As was noted before, the same kind of relationship appears between the recursion formula (4) and the original differential equation (12).

Since the solution of (12), starting from an idempotent and symmetric initial matrix, converges to a matrix satisfying both these properties, it might be asked whether this behavior of (12) is carried over the (25) in an analogous fashion. That this is indeed the case is shown by the following:

COROLLARY 1. *Let the functions* $P(\cdot)$, $Q(\cdot)$, *and* $R(\cdot)$ *as well as vector a and matrix S be as in Theorem* 2. *If* $\psi(0)$ *is idempotent and self-adjoint with respect to the inner product* $(\cdot\,,\cdot)_S$, *then, likewise, is* $\lim_{t\to\infty} \psi(t)$.

Proof. From the definition of the inner product $(\cdot\,,\cdot)_S$ in (5), it is evident that $\psi(0)$ is self-adjoint with respect to this inner product if and only if $S\psi(0)$ is symmetric. Thus $\psi(0)$ satisfies exactly the conditions of Theorem 2, implying that $\lim_{t\to\infty} \psi(t)$ is of the form (30). Squaring the limit matrix of (30) shows that it is idempotent. That it is also self-adjoint with respect to $(\cdot\,,\cdot)_S$, or $S \lim_{t\to\infty} \psi(t)$ is symmetric, follows from the symmetricity of $S\psi(0)$.

A matrix that is idempotent but not symmetric is a nonorthogonal projection matrix [7]. Thus the analogy between the behavior of $\phi(t)$ of (12) and that of $\psi(t)$ of (25) is complete: Equation (12), turning orthogonal projection matrices into orthogonal projection matrices by integration over an infinite interval, may be regarded as a special case of (25), the role of S-orthogonal projections being taken by ordinary orthogonal projections. The same analogy is, of course, revealed in the two Gram–Schmidt procedures of (4) and (10).

REFERENCES

[1] A. ALBERT, *Regression and the Moore–Penrose Pseudoinverse*, Academic Press, New York, 1972.

[2] J. K. HALE, *Ordinary Differential Equations*, John Wiley, New York, 1969.

[3] T. KOHONEN, *Associative Memory—A System-theoretical Approach*, Springer-Verlag, New York, Heidelberg and Berlin, 1977.

[4] ———, Private communication.

[5] T. KOHONEN AND E. OJA, *Fast adaptive formation of orthogonalizing filters and associative memory in recurrent networks of neuron-like elements*, Biol. Cybernet., 21 (1976), pp. 85–95.

[6] E. OJA, *Asymptotic solutions of a class of matrix differential equations arising in neural network modelling*, Internat. J. Systems Sci., 8 (1977), pp. 1145–1161.

[7] C. R. RAO AND S. K. MITRA, *Generalized Inverse of Matrices and its Applications*, John Wiley, New York, 1971.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0010 $01.00/0

AN ASYMPTOTIC PROBLEM FOR A POSITIVE DEFINITE OPERATOR-VALUED VOLTERRA KERNEL*

OLOF J. STAFFANS†

Abstract. We study the asymptotic behavior of the bounded solutions of a Volterra integral equation in a Hilbert space. The equation is supposed to have a D-positive definite convolution kernel and a nonlinearity which is the (possibly multivalued) gradient of a nonnegative function. The results we obtain generalize earlier finite dimensional results.

1. Introduction. We study the asymptotic properties of the abstract nonlinear Volterra equation

$$x(t)+\int_0^t A(t-s)g(x(s))\,ds=f(t) \qquad (t\in R^+). \tag{1.1}$$

Here f maps $R^+=[0,\infty)$ into a real Hilbert space H, g maps (a subset of) H into itself, and A is a function from R^+ into $L(H)$, the set of bounded linear operators on H. The solution x of (1.1) maps R^+ into H.

As a part of the definition of what it means for a function x to be a solution of (1.1) we suppose $g\circ x\in L^2_{\text{loc}}(R^+;H)$. We want to be able to differentiate (1.1), and for this reason A and f should be sufficiently smooth. More precisely, we suppose that f is locally differentiable in the L^2-sense, and that A is locally of bounded variation. The (distribution) derivative of A is then a dominated, operator-valued measure (call this measure μ), and (1.1) can be differentiated (cf. § 6 below):

$$x'(t)+\int_{[0,t]} d\mu(s)g(x(t-s))=f'(t) \quad \text{a.e. on } R^+. \tag{1.2}$$

When A is a smooth function, then (1.2) becomes

$$x'(t)+A(0)g(x(t))+\int_0^t A'(s)g(x(t-s))\,ds=f'(t) \quad \text{a.e. on } R^+, \tag{1.3}$$

i.e. the measure μ is the sum of the point mass $A(0)$ at zero and an absolutely continuous part.

We do not go into the question of existence of solutions of (1.1) (i.e. solutions satisfying $g\circ x\in L^2_{\text{loc}}(R^+;H)$), studied e.g. in [1], [2], [8] and [11]. However, it should be pointed out that it is not known whether a solution of (1.1) exists under our hypothesis. The existence proofs in general require at least $A(0)$ to be strictly positive (definite), A to be sufficiently smooth, and g to be maximal monotone.

Here we concentrate on studying the asymptotic behavior of solutions of (1.1), or equivalently, of (1.2), which are bounded in the sense that $\sup_{t\in R^+}\int_t^{t+1}|g(x(s))|^2\,ds<\infty$. We suppose that μ is positive definite (some authors refer to this by saying that A is *D-positive definite*). The key condition on g (see (6.5), (6.9) below) essentially amounts to assuming that g is the gradient (or subgradient) of a nonnegative function. The derivative f' of f should be sufficiently small at infinity.

The approach we use is the same as in [16]–[20], which treat the corresponding classical Volterra equation, and it is a development of an idea of Halanay [6]. To avoid unnecessary repetition we assume a certain familiarity with [16]–[20].

* Received by the editors January 17, 1977.

† Institute of Mathematics, Helsinki University of Technology, SF-02150 Espoo 15, Finland.

Before we can state the conclusion of Theorem 6.1 below on the asymptotic properties of (1.1) we have to generalize several fundamental concepts used in [17]–[20] to infinite dimensional spaces. The first of these concepts is the notion of the *limit set* $\Gamma(\varphi)$ of a function φ, studied in § 3. We let φ belong to $L^2_\infty(R; H)$, i.e. the space of functions $\varphi \in L^2_{\text{loc}}(R; H)$ such that $\sup_{t\in R} \int_t^{t+1} |\varphi(s)|^2\, ds < \infty$. This space is the dual of a Banach space, so we can define

$$\Gamma(\varphi) = \{\psi \in L^2_\infty(R; H) | \tau_{t_k}\varphi \to \psi \text{ weak* for some sequence } t_k \to \infty\},$$

where τ_h denotes the translation operator $\tau_h\varphi(t) = \varphi(t+h)$. It turns out that $\Gamma(\varphi)$, topologized with the weak*-topology, has the same nice properties as in the case $H = C$.

The second thing we must do is to develop the theory of positive definite measures in H. We give a necessary and sufficient condition (Theorem 4.1) for an operator-valued dominated measure to be positive definite, generalizing thereby the quite strong sufficient conditions in [10] and [12].

After that we define the spectral set $z(\mu)$ of a positive definite measure. Loosely speaking, it is the complement of the set of points where the symmetric part of the Fourier transform of μ is bounded away from zero in sufficiently many directions. This set plays a crucial role in our next problem: What does the boundedness of a particular quadratic integral with kernel μ (see (2.4) and (5.2) below) imply about the asymptotic behavior of the integrated function φ? We suppose that $\varphi \in L^2_\infty(R; H)$ so that one can define $\Gamma(\varphi)$ as above. For each $\psi \in \Gamma(\varphi)$ we define $\sigma(\psi)$ as the support of the distribution Fourier transform of ψ, and let $\sigma(\Gamma(\varphi))$ be the closure of the union of the sets $\sigma(\psi)$ as ψ varies over $\Gamma(\varphi)$. The conclusion of Theorem 5.1 below is the same as in the scalar case, namely $\sigma(\Gamma(\varphi)) \subset z(\mu)$.

In § 6 we apply the theory of the preceding section to get an asymptotic result for (1.1). The conclusion of Theorem 6.1 reads $\sigma(\Gamma(g \circ x)) \subset z(\mu)$. In particular, if $z(\mu) = \varnothing$, then we get convergence of $g(x(t))$ to zero as $t \to \infty$, at least in a weak sense. We conclude § 6 with a short discussion on how our work relates to [8], and we refer the reader to [8] and to [11, § 5] for discussions of earlier asymptotic results for (1.1).

2. Preliminaries. Basically we use the same notations and conventions as in [17]–[20], and we refer the reader to [20] for explanations of notations not defined here.

The technique we use requires complex scalars, so we imbed the Hilbert space H in its complexification (which we for simplicity also call H).

As before we let $F(X; Y)$ stand for the space of functions of type F mapping X into Y. For shortness we write $F(0, T; Y)$ instead of $F([0, T]; Y)$ whenever the set X is the interval $[0, T]$.

We use two new function spaces, namely L^2_1 and L^2_∞, which are mixed L^p-spaces. The functions φ in $L^2_1(R; H)$ and $L^2_\infty(R; H)$ are those which are locally L^2, and satisfy

$$\sum_{n\in Z} \left[\int_n^{n+1} |\varphi(s)|^2\, ds\right]^{1/2} < \infty$$

and

$$\sup_{n\in Z} \left[\int_n^{n+1} |\varphi(s)|^2\, ds\right]^{1/2} < \infty,$$

respectively. They are Banach spaces, with the norms indicated above, and L^2_∞ is the dual of L^2_1 (see [7, Thms. 1–2]).

By H_w we mean H equipped with its weak topology. In particular, $C(R; H_w)$ stands for weakly continuous functions from R into H.

The Banach space of bounded linear operators on H is denoted $L(H)$. We make an extensive use of the set $\mathrm{DM}(R^+; L(H))$ of $L(H)$-valued, dominated (Radon) measures on R^+. The basic facts about such measures are found in [5]. An $L(H)$-valued measure μ on R is (induced by) a continuous map from the set $\mathcal{K}(R)$ of continuous, scalar-valued functions with compact support into $L(H)$. We say that a measure μ on R is a measure on R^+ if it vanishes on $(-\infty, 0)$. The measure μ is dominated if there exists a positive measure ν such that

$$\left\|\int_R \varphi(s)\,d\mu(s)\right\| \leqq \int_R |\varphi(s)|\,d\nu(s) \qquad (\varphi \in \mathcal{K}(R)). \tag{2.1}$$

The smallest measure ν for which (2.1) holds is called the total variation (measure) of μ, and it is denoted $|\mu|$. If $|\mu|$ is bounded, i.e., $|\mu|(R) < \infty$, then we call μ a bounded measure. We write BM for the class of bounded measures.

The definition of an $L(H)$-valued function of bounded variation is completely analogous to the definition of a scalar-valued function of bounded variation. We say that a function of bounded variation is normalized if it is continuous from the right, and we denote the class of $L(H)$-valued, normalized functions which are locally of bounded variation and vanish on $(-\infty, 0)$ by $\mathrm{NBV}_{\mathrm{loc}}(R^+; L(H))$. If $A \in \mathrm{NBV}_{\mathrm{loc}}(R^+; L(H))$, then $T_A(t)$ is by definition the total variation of A on $(-\infty, t]$. Note that $T_A(t)$ vanishes on $(-\infty, 0)$.

There is a one-to-one correspondence between the classes $\mathrm{DM}(R^+; L(H))$ and $\mathrm{NBV}_{\mathrm{loc}}(R^+; L(H))$:

LEMMA 2.1. (i) *If* $\mu \in \mathrm{DM}(R^+; L(H))$, *and if*

$$A(t) = \mu((-\infty, t]) \qquad (t \in R), \tag{2.2}$$

then $A \in \mathrm{NBV}_{\mathrm{loc}}(R^+; L(H))$.

(ii) *Conversely, to every* $A \in \mathrm{NBV}_{\mathrm{loc}}(R^+; L(H))$ *there corresponds a unique measure* $\mu \in \mathrm{DM}(R^+; L(H))$ *such that* (2.2) *holds*; *for this* μ, $T_A(t) = |\mu|((-\infty, t])$ $(t \in R)$.

We only outline the proof of Lemma 2.1, which is not difficult. The proof of (i) is completely analogous to the proof of the corresponding scalar statement. To prove the opposite direction one first constructs an obvious map from the set of step functions which are continuous from the left and have compact support into $L(H)$ (cf. [13, pp. 225–226]). This map is then extended to $\mathcal{K}(R)$ by continuity, and it is easy to see that the measure μ which one gets in this way is dominated by the total variation measure of A. (This part of the proof essentially only amounts to defining the Riemann–Stieltjes integral of a continuous function.)

The convolution

$$\mu * \varphi(t) = \int_{[0,t]} d\mu(s)\varphi(t-s) \tag{2.3}$$

of a measure $\mu \in \mathrm{DM}(R^+; L(H))$ and a function $\varphi \in L^2_{\mathrm{loc}}(R^+; H)$ is well defined. More specifically, we claim that the integral in (2.3) is defined a.e. $[m]$ (m denotes the Lebesgue measure), i.e. the set of points $t \in R$ for which the function $s \to \varphi(t-s)$ is not μ-integrable has m-measure zero, and that the map $\varphi \to \mu * \varphi$ is continuous from $L^2(0, T; H)$ into itself for each $T \in R^+$. It suffices to show this in the case when $\mu \in \mathrm{BM}(R^+; L(H))$ and $\varphi \in L^2(R^+; H)$, because for each $T \in R^+$, $\chi_{[0,T]}\mu \in$

$\mathrm{BM}(R^+; L(H))$, $\chi_{[0,T]}\varphi \in L^2(R^+; H)$ and

$$\mu * \varphi(t) = (\chi_{[0,T]}\mu) * (\chi_{[0,T]}\varphi)(t) \qquad (t \in [0, T]).$$

This particular case becomes a part of [5, Prop. 24.22], if one uses the fact that R is countable at infinity to remove one assumption in that proposition:

LEMMA 2.2. *Let $\mu \in \mathrm{DM}(R; L(H))$, and let φ be an H-valued, m-measurable function on R. Then the map $(s, t) \to \varphi(t-s)$ is $|\mu| \otimes m$-measurable.*

Here $|\mu| \otimes m$ is the (completed) product of the measures $|\mu|$ and m.

Proof. There exists a Borel function $\psi \doteq \varphi$ a.e. $[m]$. As the function $(s, t) \to \psi(t-s)$ is $|\mu| \otimes m$-measurable, it suffices to show that the set $A = \{(s, t) | \psi(t-s) \neq \varphi(t-s)\}$ has outer $|\mu| \otimes m$-measure zero. Pick some Borel set E such that $m(E) = 0$, and $\psi(t) = \varphi(t)$ $(t \notin E)$. Define $F = \{(s, t) | t - s \in E\}$. Then $A \subset F$, and F is $|\mu| \otimes m$-measurable, so it suffices to show that $|\mu| \otimes m(F) = 0$. This is equivalent to

$$\int \chi_F(s, t)\, d|\mu| \otimes m(s, t) = \int \chi_E(t-s)\, d|\mu| \otimes m(s, t) = 0.$$

However, Fubini's theorem yields

$$\int \chi_E(t-s)\, d|\mu| \otimes m(s, t) = \int \left[\int \chi_E(t-s)\, dt\right] d|\mu|(s) = 0,$$

so the proof is complete.

To avoid an extensive use of brackets representing inner products we have found it quite convenient to use the operator * in the same way as in [20]. As before, for $A \in L(H)$, A^* is the adjoint of A. For $\alpha \in H$ we define α^* as the linear functional $\alpha^*\beta = \langle \beta, \alpha \rangle$ $(\beta \in H)$, where $\langle \cdot, \cdot \rangle$ is the inner product in H. Clearly, $\alpha\alpha^*$ is the $|\alpha|^2$-multiple of the orthogonal projection onto the one-dimensional subspace spanned by α. The same definitions carry over to vector- and operator-valued functions.

For $\mu \in \mathrm{DM}(R^+; L(H))$, $\varphi \in L^2(0, T; H)$ we define the quadratic form

$$Q(\mu, \varphi, T) = \mathrm{Re} \int_{[0,T]} \varphi^*(t) \int_{[0,t]} d\mu(s) \varphi(t-s)\, dt. \tag{2.4}$$

By the preceding remarks on (2.3) this definition makes sense.

3. The limit set. As in [18] we use the notation $S(\varphi)$ for the curve

$$S(\varphi) = \{\tau_t\varphi | t \in R^+\},$$

where φ is an arbitrary function in $L^2_\infty(R; H)$. By $\bar{S}(\varphi)$ we mean the weak*-closure in $L^2_\infty(R; H)$ of $S(\varphi)$. These notations remain fixed throughout this section.

In the discussion of $\bar{S}(\varphi)$ and $\Gamma(\varphi)$ (defined in § 1) we suppose throughout that H is separable. This is no loss of generality, because, as the function φ is strongly measurable, the values of φ (redefined on a set of measure zero) lie in a separable Hilbert space H_0 (see e.g. [5, Thm. 10.5]), and one can throughout replace H by H_0.

LEMMA 3.1. *Let $\varphi \in L^2_\infty(R; H)$. Give $\bar{S}(\varphi)$ the induced weak*-topology. Then $\bar{S}(\varphi)$ is compact, connected and metrizable. The limit set $\Gamma(\varphi)$ is a nonempty, closed and connected subset of $\bar{S}(\varphi)$ (hence compact), and the distance from $\tau_t\varphi$ to $\Gamma(\varphi)$ tends to zero as $t \to \infty$.*

Proof. The proof of [18, Lemma 2.1] applies with more or less trivial modifications. In particular, the separability of H implies that $L^2_1(R; H)$ is separable, and thus the weak*-topology is metrizable on $\bar{S}(\varphi)$. It is also true that translation is a continuous operation in $L_1^2(R; H)$ (cf. [7, Thm. 7]), hence a weak*-continuous operation $L_\infty^2(R; H)$.

Compared to [18, Lemma 2.1] we have not only replaced C by H, but also weakened the assumption on the local behavior of φ, replacing L^∞ by L^2. This is important in the study of the asymptotic behavior of (1.1), because it matches exactly the smoothness which some proofs of existence of solutions give.

As in [18], if the function φ belongs to some appropriate subspace of L_∞^2, then one can replace the weak*-convergence in L_∞^2 by a stronger convergence. In all cases the central argument remains the same:

LEMMA 3.2. *Let Φ be a compact space consisting of H-valued functions on R. Let $S(\varphi) \subset \Phi$, and let Φ be weak*-continuously imbedded in $L_\infty^2(R; H)$. Then $\bar{S}(\varphi) \subset \Phi$, and the topology of Φ is equivalent to the induced weak*-topology of $L_\infty^2(R; H)$.*

Proof. The compactness of Φ together with the continuous imbedding implies that Φ is weak*-compact in $L_\infty^2(R; H)$ hence weak*-closed. Thus in particular, $\bar{S}(\varphi) \subset \Phi$. The rest of the assertion follows from [14, p. 61].

In the special case when $\varphi \in L^\infty(R; H)$ we get

LEMMA 3.3. *Let $\varphi \in L^\infty(R; H)$. Then $\bar{S}(\varphi) \subset L^\infty(R; H)$, and on $\bar{S}(\varphi)$ the induced weak*-topologies of L^∞ and L_∞^2 are equivalent.*

Proof. Let Φ be the set $\{\psi \in L^\infty(R; H) \mid \|\psi\|_\infty \leqq \|\varphi\|_\infty\}$ equipped with the weak* topology of L^∞ (L^∞ is the dual of L^1), and apply Lemma 3.2.

One cannot expect to be able to define $\Gamma(\varphi)$ in terms of pointwise convergence unless φ satisfies a Tauberian condition. The classical condition when $H = C$ is

$$\lim_{t\to\infty, s\to 0} \{\varphi(t+s) - \varphi(t)\} = 0.$$

Depending on whether we put the weak or the strong topology on H we get two possible generalizations:

$$\lim_{t\to\infty, s\to 0} \{\varphi(t+s) - \varphi(t)\} = 0 \quad \text{in } H_w, \tag{3.1}$$

$$\lim_{t\to\infty, s\to 0} \{\varphi(t+s) - \varphi(t)\} = 0 \quad \text{in } H. \tag{3.2}$$

Any function which has a weak or a strong limit at ∞ satisfies (3.1) or (3.2), and so does any weakly or strongly uniformly continuous function. Actually these two classes of functions generate the whole class of functions satisfying (3.1) or (3.2) in the following sense: Every φ satisfying (3.1) or (3.2) can be split into the sum of two functions x and y, with x weakly or strongly uniformly continuous, and y tending weakly or strongly to zero (see Appendix). Clearly, if φ is split in this way, then $\Gamma(\varphi) = \Gamma(x)$, and $\tau_t\varphi$ tends pointwise to a function in $\Gamma(\varphi)$ iff $\tau_t x$ tends pointwise to the same function. For this reason one can instead of assuming (3.1) or (3.2) just as well take φ to be weakly or strongly uniformly continuous.

We begin with the weakly uniformly continuous case. Pick a countable dense subset A of H, and define $\mathscr{T}$ to be the vector space topology on $C(R; H_w)$ induced by the seminorms

$$p_{n,\alpha}(\varphi) = \sup_{t\in[-n,n]} |\alpha^*\varphi(t)| \qquad (n \in N, \alpha \in A) \tag{3.3}$$

[14, Thm. 1.37]. This topology is metrizable [14, Remark 1.38(c)], so in particular sequential compactness is equivalent to compactness. For $\varphi \in \mathrm{BUC}\,(R; H_w)$ (bounded and uniformly continuous), consider the set

$$\Phi = \{\psi \in \mathrm{BUC}\,(R; H_w) \mid \|\psi\|_\infty \leqq \|\varphi\|_\infty, \textit{ and for each } \alpha \in A \textit{ the modulus of continuity of } \alpha^*\psi \textit{ is dominated by the modulus of continuity of } \alpha^*\varphi\},$$

topologized with the induced $\mathcal{T}$-topology. It follows from Arzela–Ascoli's theorem (see e.g. [13, p. 179] with $Y = C$) together with a diagonalization process that Φ is sequentially compact, hence compact. The set S of step functions with compact support and values in A is dense in $L_1^2(R; H)$ (argue as in [5, p. 147]), and clearly $\mathcal{T}$-convergence of ψ_n to ψ in Φ implies

$$\int_R s^*(t)(\psi_n(t) - \psi(t))\, dt \to 0 \qquad (s \in S).$$

This, together with the fact that Φ is bounded in L_∞^2, yields that the imbedding of Φ in $L_\infty^2(R; H)$ is weak*-continuous. One can now apply Lemma 3.2, and one concludes that $\mathcal{T}$ coincides with the induced weak*-topology of L_∞^2.

There is yet one question which has to be resolved: Are these two topologies also equivalent on Φ to the topology $\mathcal{U}$ of weak uniform convergence on compact sets induced by the (uncountable) family of seminorms

$$p_{n,\beta}(\varphi) = \sup_{t \in [-n,n]} |\beta^*\varphi(t)| \qquad (n \in N, \beta \in H)? \tag{3.4}$$

We claim that this is the case. Clearly $\mathcal{T} \subset \mathcal{U}$, so it suffices to show that $\mathcal{U} \subset \mathcal{T}$. This follows if one can prove that for each $\varphi_0 \in \Phi$ and each $p_{n,\beta}$ in (3.4) the function $f(\varphi) = p_{n,\beta}(\varphi - \varphi_0)$ is $\mathcal{T}$-continuous on Φ. Fix f as above, and take a sequence $\alpha_m \in A$ converging strongly to β. Each function $f_m(\varphi) = p_{n,\alpha_m}(\varphi - \varphi_0)$ is $\mathcal{T}$-continuous, and $f_m \to f$ uniformly on Φ as $m \to \infty$. This implies the $\mathcal{T}$-continuity of f.

Summarizing the preceding argument we get

LEMMA 3.4. *Let $\varphi \in \mathrm{BUC}\,(R; H_w)$. Then $\bar{S}(\varphi) \subset \mathrm{BUC}\,(R; H_w)$, and on $\bar{S}(\varphi)$ the weak*-topology of $L_\infty^2(R: H)$ is equivalent to the topology of weak uniform convergence on compact sets.*

The strongly uniformly continuous case is easier to handle than the weakly uniformly continuous one:

LEMMA 3.5. *Let $\varphi \in \mathrm{BUC}\,(R; H)$, and suppose that the image of φ in H is relatively compact. Then $\bar{S}(\varphi) \in \mathrm{BUC}\,(R; H)$, and on $\bar{S}(\varphi)$ the weak*-topology of $L_\infty^2(R; H)$ is equivalent to the topology of uniform convergence on compact sets.*

Note in particular the requirement that the image of φ in H is relatively compact, which has no direct counterpart in Lemma 3.3–3.4.

Proof. Define the topology on $C(R; H)$ of uniform convergence on compact sets as in [14, Example 1.44], replacing C by H. This is a metrizable topology. Let Φ be the set

$$\Phi = \{\psi \in \mathrm{BUC}\,(R; H) \mid \textit{the image of } \psi \textit{ is contained in the closure of the image of } \varphi, \textit{ and the modulus of continuity of } \psi \textit{ is dominated by the modulus of continuity of } \varphi\}.$$

This set is compact by Arzela–Ascoli's theorem [13, p. 179]. Lemma 3.2 yields the desired conclusion.

The preceding list of different convergence concepts is by no means exhausting. One could continue e.g. by requiring φ to be differentiable, putting boundedness conditions on various derivatives, and showing that Lemma 3.2 can be applied. One could, on the other hand, go in the other direction and use topologies weaker than the weak*-topology of $L^2_\infty(R; H)$. This makes it possible to study $\bar{S}(\varphi)$ for functions φ not in $L^2_\infty(R; H)$. However, the cases we have considered here cover our present needs.

We conclude this section with a lemma on the relationship between different limit sets:

LEMMA 3.6. *Let A be a weak*-continuous map from $\bar{S}(\varphi)$ into $L^2_\infty(R^+; H_1)$* (where H_1 is some Hilbert space) *which commutes with translation. Then* $\Gamma(A\varphi) = A\Gamma(\varphi)$.

Proof. By the continuity of A, if $\tau_{t_k}\varphi \to \psi \in \Gamma(\varphi)$, then $\tau_{t_k}A\varphi = A\tau_{t_k}\varphi \to A\psi \in \Gamma(A\varphi)$. Thus $A\Gamma(\varphi) \subset \Gamma(A\varphi)$.

Conversely, let $\tau_{t_k}A\varphi \to \xi \in \Gamma(A\varphi)$. By the sequential compactness of $\bar{S}(\varphi)$ we can find a subsequence s_k of t_k such that $\tau_{s_k}\varphi \to \psi \in \Gamma(\varphi)$. The first part of the proof yields $\xi = A\psi$. Thus $\Gamma(A\varphi) \subset A\Gamma(\varphi)$.

4. Positive definite measures. In [20] we developed a theory for positive definite operator-valued measures in finite dimensional spaces. Part of that theory carries over to infinite dimensions. The definition of positive definiteness is straightforward for operator-valued dominated measures ($Q(\mu, \varphi, T)$ is defined in (2.4)):

DEFINITION 4.1. A measure $\mu \in \mathrm{DM}(R^+; L(H))$ is *positive definite* ($\mu \in \mathrm{PD}(R^+; L(H))$ if for every $T \in R^+$ and every $\varphi \in L^2(0, T; H)$,

$$Q(\mu, \varphi, T) \geqq 0. \tag{4.1}$$

Every weakly continuous operator-valued positive definite function S on R (see [21, p. 25]) induces a positive definite measure, i.e. the measure $d\mu(t) = S(t)\,dt$ ($t \in R^+$) is positive definite. The proof of this fact, outlined below, is the same as in the scalar case. It suffices to verify (4.1) for continuous functions φ, because this class is dense in $L^2(0, T; H)$, and $Q(\mu, \varphi, T)$ depends continuously on φ in the $L^2(0, T; H)$-norm. For a continuous φ the function $\varphi^*(t)S(t-s)\varphi(s)$ is continuous in s, t. Hence $Q(\mu, \varphi, T)$ can be approximated by a Riemann sum, which is nonnegative by the positive definiteness of S.

In particular we observe that every strongly continuous contraction semigroup of bounded operators on H induces a positive definite measure [21, pp. 29–30].

In [20] we gave a necessary and sufficient Fourier transform condition for positive definiteness in the case $H = C^n$. This condition does not automatically generalize to infinite dimensions. What makes the situation more complicated is that positive definiteness of μ does not in general imply that $\hat{\mu} + \hat{\mu}^*$ is a dominated measure. One can of course simply suppose that $\hat{\mu} + \hat{\mu}^*$ is well defined and a dominated measure, or put additional conditions on μ which imply that this is the case, but that leads to a certain loss of generality. We have chosen to approach the problem in a slightly different way, which involves no loss of generality: We base our theory on the equivalence (i)$\Leftrightarrow$(ii) of [20, Corollary 3.1], which is valid also here:

THEOREM 4.1. *Let $\mu \in \mathrm{DM}(R^+; L(H))$, where H is a complex Hilbert space. Then the following two statements are equivalent*:

(i) $\mu \in \mathrm{PD}(R^+; L(H))$,

(ii) $\alpha^*\mu\alpha \in \mathrm{PD}(R^+; C)$ *for every α in H.*

Proof. Trivially (i)$\Rightarrow$(ii). Suppose that (ii) holds. We first verify (4.1) in the case when the range of φ lies in some finite dimensional subspace E of H. Let P be the orthogonal projection of H onto E. Then $Q(\mu, \varphi, T) = Q(P\mu, \varphi, T)$. By (ii) and [20, Corollary 3.1], the measure $P\mu|_E$ is positive definite in E, and consequently we get (4.1) in this special case. The set of functions $\varphi \in L^2(0, T; H)$ with finite dimensional range is dense in $L^2(0, T; H)$ and $Q(\mu, \varphi, T)$ depends continuously on φ in the $L^2(0, T; H)$-norm. Thus we get (4.1) for all $\varphi \in L^2(0, T; H)$, and the proof is complete.

Theorem 4.1 is not true for real Hilbert spaces; a counterexample is given in [20, Remark 3.2]. However, after a suitable modification one can apply Theorem 4.1 also in a real Hilbert space H_R. Let H_C be the complexification of H_R ($H_C = H_R + iH_R$). Every operator $A \in L(H_R)$ extends to an operator in $L(H_C)$ ($A(\alpha + i\beta) = A\alpha + iA\beta$), and in the same way a measure $\mu \in \mathrm{DM}(R^+; L(H_R))$ extends to a measure in $\mathrm{DM}(R^+; L(H_C))$. The original measure is positive definite over H_R iff the extended measure is positive definite over H_C (simply split $\varphi \in L^2(0, T; H_C)$ into its real and imaginary parts). Thus one can apply Theorem 4.1 in H_R, provided it is modified so that α in (ii) takes it values in H_C and not in H_R. In the special case when μ is selfadjoint it suffices to take $\alpha \in H_R$ in (ii), because then for any two vectors $\alpha, \beta \in H_R$,

$$(\alpha + i\beta)^* \mu (\alpha + i\beta) = \alpha^* \mu \alpha + \beta^* \mu \beta,$$

so the positive definiteness of $\alpha^* \mu \alpha$ for $\alpha \in H_C$ follows from the positive definiteness of $\alpha^* \mu \alpha$ for $\alpha \in H_R$.

COROLLARY 4.1. *Let $a \in L^1_{\mathrm{loc}}(R^+; L(H))$, and suppose that for each $\alpha \in H$, $\alpha^* a \alpha$ is nonnegative, nonincreasing and convex. Then a defines a positive definite measure.*

Proof. Combine Theorem 4.1 with Theorem 2.3 of [17].

This corollary generalizes Corollary (4.1) of [12].

COROLLARY 4.2. *Let $\mu \in \mathrm{DM}(R^+; L(H))$ and suppose that for each $\alpha \in H$, $\alpha^* \mu([0, t])\alpha$ is a nonnegative and nonincreasing function of t on R^+. Then μ is positive definite.*

Proof. Apply Theorem 4.1 and Proposition 7.2 of [20] (with $n = 1$).

Corollary 4.2 generalizes the result one gets by combining Theorems 2.1 and 2.2 of [10].

COROLLARY 4.3. *If $\mu \in \mathrm{BM}(R^+; L(H))$, then $\mu \in \mathrm{PD}(R^+; L(H))$ iff $\hat{\mu}(\omega) + \hat{\mu}^*(\omega) \geqq 0$ ($\omega \in R$).*

Here we define $\hat{\mu}(\omega) = \int_{R^+} e^{-i\omega s}\, d\mu(s)$ ($\omega \in R$), and let $\hat{\mu}(\omega) + \hat{\mu}^*(\omega) \geqq 0$ mean $\alpha^*(\hat{\mu}(\omega) + \hat{\mu}^*(\omega))\alpha \geqq 0$ ($\alpha \in H$). Corollary 4.3 follows immediately from Theorem 4.1, because $\alpha^*(\hat{\mu} + \hat{\mu}^*)\alpha = 2 \operatorname{Re} \alpha^* \hat{\mu} \alpha = 2 \operatorname{Re} (\alpha^* \mu \alpha)\hat{}$ ($\alpha \in H$).

5. Asymptotic theory. Our next goal is to develop an asymptotic theory similar to the one in [20, § 4]. Throughout in this section we suppose (Q is defined in (2.4))

$$\mu \in \mathrm{PD}(R^+; L(H)), \qquad \varphi \in L^2_\infty(R; H), \tag{5.1}$$

$$\sup_{T \in R^+} Q(\mu, \varphi, T) < \infty. \tag{5.2}$$

DEFINITION 5.1. The *spectrum* $\sigma(\psi)$ of $\psi \in L^2_\infty(R; H)$ is the support of the distribution Fourier transform $\hat{\psi}$ (cf. [15, Chap. 1, pp. 61, 73]. The *spectrum* $\sigma(\Gamma(\varphi))$ is the closure of $\bigcup_{\psi \in \Gamma(\varphi)} \sigma(\psi)$.

We want to obtain an inclusion of the form $\sigma(\Gamma(\varphi)) \subset Z(\mu)$, where $Z(\mu)$ is "the spectral set" of μ. A part of the problem is to find a good definition of $Z(\mu)$. Clearly, the smaller we can make $Z(\mu)$, the stronger the inclusion $\sigma(\Gamma(\varphi)) \subset Z(\mu)$ will be. The method used in [20] suggests:

DEFINITION 5.2. The *regular set* $\Omega(\mu)$ is the set of points $\omega \in R$ for which there exists $\varepsilon > 0$ such that for each $\alpha \in H$ the scalar measure $d\{\operatorname{Re} (\alpha^* \mu \alpha)\hat{}\}(t) - \varepsilon |\alpha|^2\, dt$ is positive in $(\omega - \varepsilon, \omega + \varepsilon)$. Its complement is called the *spectral set* of μ, and is denoted $Z(\mu)$.

This is just a new way of writing [20, Definitions 4.1–4.2], and it makes sense even when H is infinite dimensional. If $\mu \in \mathrm{BM} \cap \mathrm{PD}\,(R^+; L(H))$, then $\omega_0 \in \Omega(\mu)$ iff there exists $\varepsilon > 0$ such that the operator-valued function $\hat{\mu}(\omega) + \hat{\mu}^*(\omega) - \varepsilon I$ is positive in $(\omega_0 - \varepsilon, \omega_0 + \varepsilon)$.

With the preceding definition of $Z(\mu)$ the argument of Theorem 4.1 of [20] goes through, and one gets indeed $\sigma(\Gamma(\varphi)) \subset Z(\mu)$. However, one does better if one from the very beginning looks in just one direction of H at a time:

DEFINITION 5.3. The *regular set* $\Omega_\alpha(\mu)$ *of* μ *in the direction* α is the set of points $\omega \in R$ for which there exists $\varepsilon > 0$ such that for each $\beta \in H$ the scalar measure $d\{\mathrm{Re}\,(\beta^* \mu \beta)\hat{}\,\}(t) - \varepsilon |\alpha^* \beta|^2\, dt$ is positive in $(\omega - \varepsilon, \omega + \varepsilon)$. Its complement is called the *spectral set of* μ *in the direction* α, and is denoted $Z_\alpha(\mu)$. For each subset A of H we let $Z_A(\mu)$ be the closure of $\bigcup_{\alpha \in A} Z_\alpha(\mu)$, and finally put $z(\mu) = \bigcap_A Z_A(\mu)$, where A varies over all separating subsets of H.

By a *separating subset* we mean one whose orthogonal complement is zero. Clearly $z(\mu) \subset Z_H(\mu) \subset Z(\mu)$. It is also clear that if μ is a scalar measure times the identity, then $z(\mu) = Z(\mu)$. We shall later give an example where $z(\mu) = \varnothing$, $Z_H(\mu) = Z(\mu) = R$.

In the case when $\mu \in \mathrm{BM} \cap \mathrm{PD}\,(R^+; L(H))$ we have $\omega \in \Omega_\alpha(\mu)$ iff there exists $\varepsilon > 0$ such that $\hat{\mu} + \hat{\mu}^* - \varepsilon \alpha \alpha^* \geqq 0$ in $(\omega - \varepsilon, \omega + \varepsilon)$.

LEMMA 5.1. *Let* (5.1)–(5.2) *hold. Then for each* $\alpha \in H$, $\sigma(\Gamma(\alpha^* \varphi)) \subset Z_\alpha(\mu)$.

Proof. Let $\omega_0 \in \Omega_\alpha(\mu)$, and let ε be as in Definition 5.3. Pick some $\eta \in \mathcal{S}(R; C)$ such that $\hat{\eta}(\omega) = 0$ $(\omega \notin (\omega_0 - \varepsilon, \omega_0 + \varepsilon))$, $0 < \hat{\eta}(\omega) < 2\varepsilon$ $(\omega \in (\omega_0 - \varepsilon, \omega_0 + \varepsilon))$, and define a as the restriction of η to R^+. With the aid of Theorem 4.1 and Corollary 1.1 of [17] we conclude that the operator-valued measure $d\mu(t) - \alpha a(t) \alpha^*\, dt$ is positive definite (note that $\mathrm{Re}\, \hat{a} = \frac{1}{2}\hat{\eta}$). Thus by (2.4) and (5.2),

$$\sup_{T \in R^+} \mathrm{Re} \left\{ \int_{[0,T]} \overline{\alpha^* \varphi(t)} \int_{[0,t]} \eta(s) \alpha^* \varphi(t-s)\, ds\, dt \right\} < \infty.$$

The desired conclusion $\omega_0 \notin \sigma(\Gamma(\alpha^* \varphi))$ would follow if we could apply Theorem 3.1 of [18] (combined with Lemma 1.1 of [17]). Formally, we cannot do it, because we have only $\alpha^* \varphi \in L^2_\infty(R; C)$ and not $\alpha^* \varphi \in L^\infty(R; C)$. However, one can use the same argument as in [16] and [18] to show that every function $\psi \in \Gamma(\alpha^* \varphi)$ satisfies $\eta * \psi = 0$. Convolving this equation with an approximate identity δ_n $(n \in N)$ (see [14, p. 157]) and using the commutativity of convolution together with Theorem 9.3 of [14] we get $(\omega_0 - \varepsilon, \omega_0 + \varepsilon) \cap \sigma(\delta_n * \psi) = \varnothing$ $(n \in N)$. But $\delta_n * \psi \to \psi$ in $\mathcal{S}'$ as $n \to \infty$, so $\omega_0 \notin \sigma(\psi)$. This completes the proof of Lemma 5.1.

THEOREM 5.1. *Let* (5.1)–(5.2) *hold. Then* $\sigma(\Gamma(\varphi)) \subset z(\mu)$.

Proof. We must show that for each separating subset A of H we have $\sigma(\Gamma(\varphi)) \subset Z_A(\mu)$. Fix A. By Lemma 5.1,

$$\bigcup_{\alpha \in A} \sigma(\Gamma(\alpha^* \varphi)) \subset Z_A(\mu).$$

Lemma 3.6 (with $H_1 = C$) yields $\Gamma(\alpha^* \varphi) = \alpha^* \Gamma(\varphi)$, so we get

$$\bigcup_{\alpha \in A} \sigma(\alpha^* \Gamma(\varphi)) \subset Z_A(\mu).$$

But this implies $\sigma(\Gamma(\varphi)) \subset Z_A(\mu)$, because A is separating and $Z_A(\mu)$ is closed (cf. [15, Chap. 1, p. 61]; two distributions u and v are equal if $\alpha^* u = \alpha^* v$ for α in a separating subset of H). This completes the proof of Theorem 5.1.

Analogues to Theorems 4.1 and 4.3 of [18] also hold.

There are cases where the set $z(\mu)$ is strictly smaller than the sets $Z_H(\mu)$ and $Z(\mu)$. This is most easily demonstrated with the following example: Let ν be a scalar, finite measure with Re $\hat{\nu}(\omega)>0$ ($\omega \in R$), let $A \in L(H)$ be positive, selfadjoint, having zero in its continuous spectrum but not in its point spectrum, and define $\mu = A\nu$. We claim that $z(\mu)=\emptyset$. Let E be the resolution of the identity associated with A [14, Thm. 13.30], and let Y_λ be the range of $E((\lambda, \infty))$. As zero is not an eigenvalue of A, the space $Y=\bigcup_{\lambda>0} Y_\lambda$ is dense in H, hence separating. We claim that $Z_Y(\mu)=\emptyset$, i.e. $Z_\alpha(\mu)=\emptyset$ ($\alpha \in Y$). Obviously this implies our earlier claim $z(\mu)=\emptyset$. Let $\alpha \in Y$. Then we can find $\lambda>0$ such that $\alpha \in Y_\lambda$. Take an arbitrary $\beta \in H$, and decompose it into $\beta_\lambda \in Y_\lambda$ and $\beta-\beta_\lambda \in Y_\lambda^\perp$. For an arbitrary $\omega \in R$ we get

$$\begin{aligned}\operatorname{Re}(\beta^*\mu\beta)\hat{}(\omega) &= \operatorname{Re}\hat{\nu}(\omega)\beta^*A\beta\\ &= \operatorname{Re}\hat{\nu}(\omega)[\beta_\lambda^*A\beta_\lambda+(\beta-\beta_\lambda)^*A(\beta-\beta_\lambda)]\\ &\geqq \lambda \operatorname{Re}\hat{\nu}(\omega)|\beta_\lambda|^2,\end{aligned}$$

where we have used the facts that Y_λ and $Y_\lambda^\perp$ are invariant subspaces of A, that A is positive on $Y_\lambda^\perp$, and that $A \geqq \lambda I$ on Y_λ. As $|\alpha^*\beta|=|\alpha^*\beta_\lambda| \leqq |\alpha|\,|\beta_\lambda|$, we find that $Z_\alpha(\mu)=\emptyset$. Consequently $Z_Y(\mu)=z(\mu)=\emptyset$.

We claim that, on the other hand, $Z_H(\mu)=R$. To prove this it suffices to find some $\alpha \in H$ such that $Z_\alpha(\mu)=R$. Take an orthogonal sequence of unit vectors $\alpha_n \in Y_{4^{-n}}^\perp$ (this is possible because zero is in the continuous spectrum of A), and define $\alpha=\sum_{n\in N} 2^{-n}\alpha_n$. Then

$$|\alpha^*\alpha_n|^{-2}\operatorname{Re}(\alpha_n^*\mu\alpha_n)\hat{}(\omega)=2^n\operatorname{Re}\hat{\nu}(\omega)\alpha_n^*A\alpha_n \leqq 2^{-n}\operatorname{Re}\hat{\nu}(\omega)\to 0 \qquad (n\to\infty),$$

which shows that $Z_\alpha(\mu)=R$.

6. An integral equation. We want to apply the preceding theory to study the asymptotic properties of the integral equation

$$x(t)+\int_0^t A(t-s)g(x(s))\,ds \ni f(t) \qquad (t \in R^+). \tag{6.1}$$

Here g is a possibly multivalued map from its domain D_g contained in H (which we here take to be a real Hilbert space) into H. We suppose that

$$A \in \mathrm{NBV}_{\mathrm{loc}}(R^+; L(H)), \tag{6.2}$$

$$f \in \mathrm{LAC}(R^+; H), \qquad f' \in L^2_{\mathrm{loc}}(R^+; H), \tag{6.3}$$

where LAC stands for the class of locally absolutely continuous functions. We call x a solution of (6.1) if it satisfies (6.4)–(6.6) below:

$$x \in C(R^+; H), \qquad x(t) \in D_g \quad \text{a.e. on } R^+, \tag{6.4}$$

(6.5) there exists a function $\varphi \in L^2_{\mathrm{loc}}(R^+; H)$ such that $\varphi(t) \in g(x(t))$ a.e. on R^+,

$$x(t)+\int_0^t A(t-s)\varphi(s)\,ds=f(t) \qquad (t \in R^+). \tag{6.6}$$

We claim that a solution x of (6.1) is necessarily locally absolutely continuous and satisfies

$$x'(t)+\int_{[0,t]} d\mu(s)\varphi(t-s)=f'(t) \quad \text{a.e. on } R^+, \tag{6.7}$$

where μ is the dominated measure corresponding to A (see Lemma 2.1). To show this one first uses Fubini's theorem, then a change of variable, and then Fubini's theorem once more to verify that

$$\int_0^t A(t-s)\varphi(s)\,ds = \int_{[0,t]} \left(\int_{[0,t-s]} d\mu(v)\right)\varphi(s)\,ds$$
$$= \int_0^t \left(\int_{[0,s]} d\mu(v)\varphi(s-v)\right) ds.$$

This together with (6.3) and (6.6) implies that x is locally absolutely continuous, and differentiating (6.6) one gets (6.7).

Before we can apply the theory of §§ 3–5 we need some further assumptions:

(6.8) $$\mu \in \mathrm{PD}\,(R^+; L(H)),$$

(6.9) there exists a nonnegative function G on $D_G \supset D_g$ such that $G(x(t)) = G(x(0)) + \int_0^t \varphi^*(s)x'(s)\,ds \quad (t \in R^+)$,

(6.10) $$f' \in L_1^2(R^+; H),$$

(6.11) $$\varphi \in L_\infty^2(R^+; H).$$

THEOREM 6.1. *Let* (6.2)–(6.6) *and* (6.8)–(6.11) *hold. Then* $\sigma(\Gamma(\varphi)) \in z(\mu)$. *In particular, if* $z(\mu) = \varnothing$, *then* $\tau_t\varphi \to 0$ *weak* in* $L_\infty^2(R; H)$ *as* $t \to \infty$.

Here φ is extended to R e.g. by zero outside R^+, and H has been imbedded in its complexification so that our previous definitions of $\sigma(\Gamma(\varphi))$ and $z(\mu)$ can be applied.

Proof. Multiply (6.7) by $\varphi^*(t)$, and integrate over $[0, T]$. This yields (5.2). Thus by Theorem 5.1, $\sigma(\Gamma(\varphi)) \subset z(\mu)$.

If $z(\mu) = \varnothing$, then $\sigma(\psi) = \varnothing$ for each $\psi \in \Gamma(\varphi)$, and thus $\Gamma(\varphi) = \{0\}$. Lemma 3.1 yields $\tau_t\varphi \to 0$ weak* in $L_\infty^2(R; H)$ as $t \to \infty$, completing the proof of Theorem 6.1.

COROLLARY 6.1. *In addition to* (6.2)–(6.6), (6.8)–(6.10) *suppose that* $\varphi \in L^\infty(R^+; H)$, *and that* $z(\mu) = \varnothing$. *Then* $\tau_t\varphi \to 0$ *weak* in* $L^\infty(R; H)$ *as* $t \to \infty$.

Proof. Combine Theorem 6.1 with Lemma 3.3.

COROLLARY 6.2. *In addition to* (6.2)–(6.6), (6.8)–(6.10) *suppose that* $\varphi \in \mathrm{BUC}\,(R^+; H_w)$, *and that* $z(\mu) = \varnothing$. *Then* $\varphi(t)$ *tends weakly to zero as* $t \to \infty$.

Proof. We extend φ to R defining $\varphi(t) = \varphi(0)$ $(t < 0)$, and then apply Theorem 6.1 and Lemma 3.4.

COROLLARY 6.3. *In addition to* (6.2)–(6.6), (6.8)–(6.10) *suppose that* $\varphi \in \mathrm{BUC}\,(R^+; H)$, *that the image of* φ *in* H *is relatively compact, and that* $z(\mu) = \varnothing$. *Then* $\varphi(t) \to 0$ $(t \to \infty)$.

Proof. We extend φ to R defining $\varphi(t) = \varphi(0)$ $(t < 0)$, and then apply Theorem 6.1 and Lemma 3.5.

Theorem 6.1 overlaps Londen's Theorem 2, Corollaries 1 and 4 of [8]. We assume much less on the kernel μ than Londen does (cf. [20, Proposition 7.2] with $n = 1$), but then we have to shift (6.11) from the conclusion to the hypothesis. Our (6.9) is somewhat stronger than the corresponding assumption in [8]. The conclusion $\sigma(\Gamma(\varphi)) \subset z(\mu)$ of Theorem 6.1 has no direct counterpart in [8], but it is related to Londen's conclusion $x' \in L^2(R^+; H)$. Under Londen's assumption on μ, $z(\mu) \subset \{0\}$, so Theorem 6.1 yields $\sigma(\Gamma(\varphi)) \subset \{0\}$, i.e. $\Gamma(\varphi)$ contains nothing but constant functions. Londen's $x' \in L^2(R^+; H)$ trivially yields $\Gamma(x') = \{0\}$. Both statements $\sigma(\Gamma(\varphi)) \subset \{0\}$ and $\Gamma(x') = \{0\}$ express the fact that the solution x of (6.1) acts more and more like a constant as $t \to \infty$. It can be shown that (for Londen's kernel) the former implies the latter.

Appendix. The following lemma was referred to in § 3:

LEMMA A.1. *Let* $\varphi \in L^\infty(R; H)$ *satisfy* (3.1) *or* (3.2). *Then* $\varphi(t) = x(t) + y(t)$ $(t \in R)$, *where x is bounded and weakly or strongly uniformly continuous, and y tends weakly or strongly to zero as* $t \to \infty$.

The proof is completely similar in the weak and the strong case, so we do not distinguish between the two.

Proof. Take some $\eta \in C(R; R)$ such that $\eta(t) = 1$ $(t \leqq 0)$, $\eta(t) > 0$ $(t > 0)$, and $\eta(t) \to 0$ $(t \to \infty)$. Define

$$x(t) = [\eta(t)]^{-1} \int_t^{t+\eta(t)} \varphi(s)\, ds = \int_0^1 \varphi(t + \eta(t)s)\, ds,$$

$$y(t) = \varphi(t) - x(t) = \int_0^1 [\varphi(t) - \varphi(t + \eta(t)s)]\, ds \qquad (t \in R).$$

Then x is continuous on R, and uniformly so on $(-\infty, 0]$. It follows from (3.1) or (3.2) that $\varphi(t) - \varphi(t + \eta(t)s) \to 0$ $(t \to \infty)$ uniformly for $s \in [0, 1]$. Hence $y(t) \to 0$ $(t \to \infty)$. Thus both φ and y satisfy (3.1) or (3.2); hence so does x. Together with the continuity this yields uniform continuity of x on R^+, and completes the proof.

REFERENCES

[1] V. BARBU, *Nonlinear Volterra equations in a Hilbert space*, this Journal, 6 (1975), pp. 728–741.

[2] ———, *On a nonlinear Volterra integral equation on a Hilbert space*, this Journal, 8 (1977), pp. 346–355.

[3] N. BOURBAKI, *Intégration Vectorielle*, Hermann, Paris, 1959.

[4] M. G. CRANDALL, S.-O. LONDEN AND J. A. NOHEL, *An abstract nonlinear Volterra integrodifferential equation*, J. Math. Anal. Appl., to appear.

[5] N. DINCULEANU, *Integration on Locally Compact Spaces*, P. Noordhoff, Leyden, The Netherlands, 1974.

[6] A. HALANAY, *On the asymptotic behavior of the solutions of an integro-differential equation*, J. Math. Anal. Appl., 10 (1965), pp. 319–324.

[7] F. HOLLAND, *Harmonic analysis on amalgams of* L^p *and* l^q, J. London Math. Soc. Ser. 2, 10 (1975), pp. 295–305.

[8] S.-O.. LONDEN, *On an integral equation in a Hilbert space*, this Journal, 8 (1977), pp. 950–970.

[9] ———, *An existence result on a Volterra equation in a Banach space*, Trans. Amer. Math. Soc., to appear.

[10] R. C. MACCAMY, *Nonlinear Volterra equations on a Hilbert space*, J. Differential Equations, 16 (1974), pp. 373–393.

[11] ———, *Stability theorems for a class of functional differential equations*, SIAM J. Appl. Math., to appear.

[12] R. C. MACCAMY AND J. S. W. WONG, *Stability theorems for some functional equations*, Trans. Amer. Math. Soc., 164 (1972), pp. 1–37.

[13] H. L. ROYDEN, *Real Analysis*, 2nd ed., Macmillan, London, 1968.

[14] W. RUDIN, *Functional Analysis*, McGraw-Hill, New York, 1973.

[15] L. SCHWARTZ, *Théorie des distributions a valeurs vectorielles*, Ann. Inst. Fourier (Grenoble), 7 (1957), pp. 1–141; 8 (1958), pp. 1–209.

[16] O. J. STAFFANS, *Nonlinear Volterra integral equations with positive definite kernels*, Proc. Amer. Math. Soc., 51 (1975), pp. 103–108.

[17] ———, *Positive definite measures with applications to a Volterra equation*, Trans. Amer. Math. Soc., 218 (1976), pp. 219–237.

[18] ———, *Tauberian theorems for a positive definite form, with applications to a Volterra equation*, Ibid., 218 (1976), pp. 239–259.

[19] ———, *On the asymptotic spectra of the bounded solutions of a nonlinear Volterra equation*, J. Differential Equations, 24 (1977), pp. 365–382.

[20] ———, *Systems of nonlinear Volterra equations with positive definite kernels*, Trans. Amer. Math. Soc., 228 (1977), to appear.

[21] B. SZ.-NAGY AND C. FOIAS, *Harmonic Analysis of Operators on Hilbert Space*, North-Holland, Amsterdam, 1970.

[22] R. MCKELVEY, *Spectral measures, generalized resolvents, and functions of positive type*, J. Math. Anal. Appl., 11 (1965), pp. 447–477.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0011 $01.00/0

ON THE ASYMPTOTIC BEHAVIOR OF FINITE ENERGY SOLUTIONS OF AN ABSTRACT INTEGRAL EQUATION*

OLOF J. STAFFANS†

Abstract. We define an energy function for a solution of the abstract nonlinear integral equation

$$x'(t)+\partial\varphi(x(t))+\int_0^t a(t-s)\,\partial\psi(x(s))\,ds \ni f(t) \qquad (t\in R^+),$$

and study the asymptotic behavior of the solutions for which the energy function is bounded. We also investigate the problem of getting an a priori bound on the energy function.

1. Introduction. We study the asymptotic behavior of the solutions of the abstract integral equation

$$\text{(E)} \qquad x'(t)+\partial\varphi(x(t))+\int_0^t a(t-s)\,\partial\psi(x(s))\,ds \ni f(t) \qquad (t\in R^+)$$

(we write R^+ for the interval $[0, \infty)$). Our setting is the same as in [3]. We have a real reflexive Banach space W continuously and densely imbedded in a real Hilbert space H. We identify H with its own dual, and by transposing the imbedding of W into H we obtain an imbedding of H into W' (the dual of W). We are given two convex lower semicontinuous proper functions $\varphi: H\to(-\infty,\infty]$ and $\psi: W\to(-\infty,\infty]$, and let $\partial\varphi$ and $\partial\psi$ be their subgradients (see e.g. [2]). Then $\partial\varphi$ maps H into itself, and $\partial\psi$ maps W into W'. We suppose that ψ has a lower semicontinuous extension to H. The functions a and f belong to $C(R^+; R)$ and $L^2_{\text{loc}}(R^+; H)$, respectively.

The existence results for solutions of (E) found in [1], [3]–[5] provide us with a notion of a solution of (E) (although none of these apply without a number of additional assumptions). We call x a solution of (E) if the following conditions are satisfied:

$$\text{(1.1i)} \qquad x\in C(R^+; W), \quad x'\in L^2_{\text{loc}}(R^+; H), \quad \psi(x(0))<\infty,$$

(1.1ii) there exists $v\in L^2_{\text{loc}}(R^+; H)$ such that $v(t)\in\partial\varphi(x(t))$ a.e. on R^+,

(1.1iii) there exists $w\in L^2_{\text{loc}}(R^+; W')$ such that $w(t)\in\partial\psi(x(t))$ a.e. on R^+,

$$\text{(1.1iv)} \qquad x'+v(t)+\int_0^t a(t-s)w(s)\,ds=f(t) \quad \text{a.e. on } R^+.$$

Here x' is the distribution derivative of x. Note that, because of (1.1iii), the integral in (1.1iv) is a continuous function with values in W'. However, it follows from (1.1i)–(1.1iv) and $f\in L^2_{\text{loc}}(R^+; H)$ that the same integral also belongs to $L^2_{\text{loc}}(R^+; H)$.

We suppose in the sequel that the function a is positive definite. In this case the following method has recently been used to study the asymptotic behavior of a solution of an equation of type (E): One takes the inner product in H of $w(t)$ and (1.1iv), and integrates over $[0, T]$. With appropriate assumptions on φ, ψ and f one

* Received by the editors September 20, 1976.
† Institute of Mathematics, Helsinki University of Technology, SF-02150 Espoo 15, Finland.

then shows that

$$\sup_{T\in R^+}\int_0^T\left(w(t),\int_0^t a(t-s)w(s)\,ds\right)dt<\infty, \tag{1.2}$$

and uses (1.2) to draw conclusions about the asymptotic behavior of w. Unfortunately this approach does not work here, because the terms that one is supposed to integrate are not integrable, unless in addition to (1.1) one has $w\in L^2_{\text{loc}}(R^+;H)$. In particular, the integral

$$\int_0^T\left(w(t),\int_0^t a(t-s)w(s)\,ds\right)dt \tag{1.3}$$

is not well defined unless $w\in L^1_{\text{loc}}(R^+;H)$.

We avoid the problem of how one should define (1.3) by simply replacing it by a different quadratic form $Q(w,t)$ (see (4.2) below), which is easy to define, and which can be used in exactly the same way as (1.3) when one wants to study how w behaves asymptotically. We define $Q(w,T)$ with the aid of Fourier transforms, and it differs from (1.3) in the case when $w\in L^1_{\text{loc}}(R^+;H)$ only by the use of a norm in W' instead of a norm in H. If we suppose that $|\cdot|_H\geqq|\cdot|_{W'}$ (which we can always do by rescaling one of the norms, if necessary), then

$$Q(w,T)\leqq\int_0^T\left(w(t),\int_0^t a(t-s)w(s)\,ds\right)dt,$$

whenever the right hand side makes sense. In particular, (1.2) implies

$$\sup_{T\in R^+}Q(w,T)<\infty. \tag{1.4}$$

As we show below, (1.4) yields almost exactly the same conclusion about the asymptotic behavior of w as (1.2) does. The only difference is that one has to work in W' instead of working in H.

The question whether all solutions of (E) under reasonable conditions on φ, ψ and f (e.g. (5.2), (5.5) and (5.6) below) satisfy (1.4) remains open. However, it is very interesting to observe that some solutions do, in particular, all those that the existence theorems in [1], [3]–[5] produce (whenever they apply). Of course, this does not exclude the possibility that there exist other solutions which do not satisfy (1.4) (even when (5.2), (5.5) and (5.6) hold).

This work overlaps Theorem 3(iii) of [3], and especially a remark made in its proof. In [3] the kernel a is nonconstant, convex and sufficiently smooth, which implies that w only can have a trivial asymptotic behavior (convergence to zero). When $w\in L^2_{\text{loc}}(R^+;H)$, then a somewhat more general, strictly positive definite kernel is allowed. The condition on f used in [3] depends on the convexity of the kernel a, and it is quite different from our (5.6). On the other hand, our (5.2) has no direct counterpart in [3] (although it is somewhat similar to the local boundedness of $\partial\psi$ required in [3]). Theorem 3(ii) in [3] yields a Tauberian condition (uniform continuity) on w, and that makes it possible to state the results of Theorem 3(iii) of [3] in terms of pointwise convergence rather than in terms of the weak*-convergence in $L^2_\infty(R;W')$ used by us.

Theorem 4 of [3] contains an asymptotic result for (E) when the kernel a is not necessarily positive definite. Instead it is assumed that $\partial\varphi$ dominates $\partial\psi$ (see [3, line (1.18)]), and this makes it possible to obtain estimates which imply $w\in L^2(R^+;H)$.

This work can be considered as a continuation of [11].

2. Notations and definitions. In §§ 2–4 we let B be a complex reflexive Banach space. We are really more interested in the case when the scalars in B are real, but the technique we use requires complex scalars. Note that it is no loss of generality to take B complex, because every real Banach space B_{R} can be regarded as the "real part" of a complex Banach space B_{C}. One simply considers $B_{\mathrm{R}} \times B_{\mathrm{R}}$ as a complex Banach space $B_{\mathrm{C}} = B_{\mathrm{R}} + iB_{\mathrm{R}}$, defining multiplication by complex scalars by $(\alpha + i\beta)(x, y) = (\alpha x - \beta y, \beta x + \alpha y)$. We refer to B_{C} as the complexification of B_{R}. The dual of B is denoted B'.

Our asymptotic theorem (Theorem 4.1 below) applies to functions φ in the mixed L^p-space $L^2_\infty(R; B)$. These functions are locally L^2, and satisfy

$$\sup_{n \in Z} \int_n^{n+1} |\varphi(s)|^2 \, ds < \infty.$$

The space $L^2_\infty(R; B)$ is the dual of a Banach space (cf. [11, § 2]), and by $L^2_\infty(R; B)_{w^*}$ we mean $L^2_\infty(R; B)$ endowed with its weak*-topology.

The limit set $\Gamma(\varphi)$ of a function $\varphi \in L^2_\infty(R; B)$ is defined by

$$\Gamma(\varphi) = \{\psi \in L^2_\infty(R; B) | \tau_{t_k}\varphi \to \psi \text{ in } L^2_\infty(R; B)_{w^*}, \text{ for some sequence } t_k \to \infty\}$$

(here and in the sequel we write τ for the translation operator $\tau_h \varphi(t) = \varphi(t+h)$). Its properties are studied in § 3.

The spectrum $\sigma(\varphi)$ of a function $\varphi \in L^2_\infty(R; B)$ is by definition the support of the distribution Fourier transform of φ (see [8]).

We denote the (distribution) Fourier transform of a function φ by $\hat{\varphi}$. If φ is integrable, then $\hat{\varphi}$ is also given by $\hat{\varphi}(\omega) = \int_R e^{-i\omega t} \varphi(t) \, dt$.

Functions defined on R^+ are extended to R by zero outside R^+.

Throughout this paper the function a is taken to be continuous and positive definite (in the sense of [9, Remark 2.2]). The set $Z(a)$, referred to in Theorem 4.1 below, is basically the zero set of the real part of the Fourier transform of a. If e.g., $a \in L^1(R^+; R)$, then $\hat{a} \in C(R; C)$, and the preceding definition makes sense as it stands. In general one does not have $a \in L^1(R^+; R)$, and then a more complicated definition is needed. By Bochner's theorem, there exists a positive finite measure λ such that

$$a(t) = \frac{1}{\pi} \int_R e^{i\omega t} \, d\lambda(\omega) \qquad (t \in R^+) \tag{2.1}$$

(the measure λ is actually the real part of the distribution Fourier transform of a). We call λ strictly positive at a point ω_0 if there exists $\varepsilon > 0$ such that the measure $d\lambda(\omega) - \varepsilon \, d\omega$ is positive in $(\omega_0 - \varepsilon, \omega_0 + \varepsilon)$ (here $d\omega$ is the Lebesgue measure). Finally, we define $Z(a)$ to be the set where λ is not strictly positive.

By $\chi_{[0,T]}$ we mean the characteristic function of the interval $[0, T]$.

3. The limit set. The following basic result on the limit set $\Gamma(\varphi)$ of a function $\varphi \in L^2_\infty(R; B)$ is a minor extension of Lemma 3.1 in [11] (there B was a Hilbert space):

LEMMA. 3.1. *Let $\varphi \in L^2_\infty(R; B)$. Define $\bar{S}(\varphi)$ as the closure of $S(\varphi) = \{\tau_t \varphi | t \in R^+\}$ in $L^2_\infty(R; B)_{w^*}$, and give it the induced topology. Then $\bar{S}(\varphi)$ is compact, connected and metrizable. The limit set $\Gamma(\varphi)$ is a nonempty, closed and connected subset of $\bar{S}(\varphi)$ (hence compact), and the distance from $\tau_t\varphi$ to $\Gamma(\varphi)$ tends to zero as $t \to \infty$.*

Proof. When B is separable, then the proof in [11] (see also [10]) applies with trivial changes. If B is not separable, then one argues as follows. The values of φ (modulo a set of measure zero) lie in a closed separable subspace V of B. As closed subspaces of reflexive Banach spaces are reflexive, one can apply Lemma 3.1 with B replaced by V. It then only remains to show that the closure $\tilde{S}(\varphi)$ of $S(\varphi)$ in $L^2_\infty(R; B)_{w^*}$ is the same as the closure $\bar{S}(\varphi)$ in $L^2_\infty(R; V)_{w^*}$, and that $L^2_\infty(R; B)_{w^*}$ and $L^2_\infty(R; V)_{w^*}$ induce the same topology on $\bar{S}(\varphi)$. It is easy to see that the topology induced by $L^2_\infty(R; B)_{w^*}$ on $\bar{S}(\varphi)$ is weaker than the $L^2_\infty(R; V)_{w^*}$-topology (use [7, Thm. 4.9] and the fact that the quotient map of B' into $B'/V^\perp$ is continuous). Thus, in particular, $\bar{S}(\varphi) \subset \tilde{S}(\varphi)$. On the other hand, $\bar{S}(\varphi)$ is compact in $L^2_\infty(R; V)_{w^*}$, hence compact in $L^2_\infty(R; B)_{w^*}$, hence closed in $L^2_\infty(R; B)_{w^*}$. Thus $\bar{S}(\varphi) = \tilde{S}(\varphi)$. The equivalence of the two topologies on $\bar{S}(\varphi)$ follows from the uniqueness of a compact Hausdorff topology [7, p. 61].

Remark 3.1. Also Lemma 3.3–3.5 in [11] are true in reflexive Banach spaces. The proofs in [11] are formulated for the case when B is a Hilbert space, but after some minor modifications they apply also in the more general case when B is a reflexive Banach space.

4. An asymptotic theorem. Before we can state our asymptotic theorem we have to define the functions P and Q. For $\varphi \in L^1(R; B)$, we define

$$P(\varphi) = \frac{1}{2\pi} \int_R |\hat{\varphi}(\omega)|^2 \, d\lambda(\omega), \tag{4.1}$$

where λ is the measure in (2.1). We claim that P is convex and continuous on $L^1(R; B)$. The convexity follows trivially from the linearity of the Fourier transform and the convexity of the square of the norm in B, whereas the continuity is an immediate consequence of the fact that convergence in $L^1(R; B)$ implies uniform convergence of Fourier transforms.

The function Q is defined by

$$Q(\varphi, T) = P(\chi_{[0,T]}\varphi) \tag{4.2}$$

for $\varphi \in L^2_{\text{loc}}(R^+; B)$ and $T \in R^+$. Clearly $\chi_{[0,T]}\varphi \in L^1(R; B)$, so the preceding definition makes sense. Note that for fixed T, $Q(\varphi, T)$ is a convex and continuous function of φ in the $L^2(0, T; B)$-norm, hence also weakly lower semicontinuous in $L^2(0, T; B)$.

THEOREM 4.1. *Let* $\varphi \in L^2_\infty(R; B)$, *and suppose that* $\sup_{T \in R^+} Q(\varphi, T) < \infty$. *Then each* $\psi \in \Gamma(\varphi)$ *satisfies* $\sigma(\psi) \subset Z(a)$.

Proof. We first claim that $\alpha \in B'$, $\psi_\alpha \in \Gamma(\alpha\varphi)$ implies $\sigma(\psi_\alpha) \subset Z(a)$. Fix $\alpha \in B'$, and define $\varphi_\alpha = \alpha\varphi$. Then $\varphi_\alpha \in L^2_\infty(R; C)$. Clearly $(\chi_{[0,T]}\varphi_\alpha)^\wedge = \alpha(\chi_{[0,T]}\varphi)^\wedge$. In particular,

$$|(\chi_{[0,T]}\varphi_\alpha)^\wedge(\omega)| \leqq |\alpha|_{B'} |(\chi_{[0,T]}\varphi)^\wedge(\omega)|_B \qquad (\omega \in R).$$

This, together with $\sup_{T \in R^+} Q(\varphi, T) < \infty$ yields

$$\sup_{T \in R^+} \int_R |(\chi_{[0,T]}\varphi_\alpha)^\wedge(\omega)|^2 \, d\lambda(\omega) < \infty.$$

It is a matter of straightforward computation (see the proof of Lemma 6.1 in [9]) to show that

$$\frac{1}{\pi} \int_R |(\chi_{[0,T]}\varphi_\alpha)^\wedge(\omega)|^2 \, d\lambda(\omega) = \int_0^T \bar{\varphi}_\alpha(t) \int_0^T a(t-s)\varphi_\alpha(s) \, ds \, dt,$$

where we have defined $a(-t) = a(t)$ $(t > 0)$. Hence

$$\sup_{T \in R^+} \int_0^T \bar{\varphi}_\alpha(t) \int_0^T a(t-s)\varphi_\alpha(s) \, ds \, dt < \infty.$$

One can now apply the scalar theory developed in [10] (see [10, Thm. 3.1] and also the last paragraph in the proof of Lemma 5.1 in [11]) to conclude that $\psi_\alpha \in \Gamma(\varphi_\alpha)$ implies $\sigma(\psi_\alpha) \subset Z(a)$. This verifies our claim.

Take some arbitrary $\psi \in \Gamma(\varphi)$. Then for each $\alpha \in B'$, one has $\alpha\psi \in \Gamma(\alpha\varphi)$ (cf. [11, Lemma 3.6]). By the preceding argument, $\sigma(\alpha\psi) \subset Z(a)$, or equivalently, $\underline{\text{supp}}\,(\alpha\psi)\hat{} = \underline{\text{supp}}\,(\alpha\hat{\psi}) \subset Z(a)$, where we use $\underline{\text{supp}}$ to denote the support of a distribution. But this implies $\underline{\text{supp}}\,\hat{\psi} \subset Z(a)$ (see [8, Chap. I, p. 61]), and completes the proof of Theorem 4.1.

5. The integral equation. We want to apply Theorem 4.1 to study the asymptotic behavior of a solution of (E). Whenever we refer to a solution of (E), we are really thinking about a specific set of functions x, v, w satisfying (1.1). Thus, whenever x is given, we also consider v and w given.

Substitute $\varphi = w$ in Theorem 4.1, and let B be the complexification of W' (in particular, the norm used in (4.1) is the norm of the complexification of W'). The hypothesis then becomes

(5.1i) $$w \in L^2_\infty(R^+;\, W'),$$

(5.1ii) $$\sup_{T \in R^+} Q(w, T) < \infty.$$

We claim that (5.1) follows from (5.2)–(5.3) below:

(5.2) ψ is nonnegative, and there exists a constant K such that $z \in \partial\psi(x)$ implies $|z|_{W'} \leqq K(1+\psi(x))$.

(5.3) $$\sup_{T \in R^+} E(T) < \infty, \quad \text{where } E(T) = \psi(x(T)) + Q(w, T).$$

Since both $\psi(x(T))$ and $Q(w, T)$ are nonnegative, the condition (5.3) implies both (5.1ii) and

(5.4) $$\sup_{T \in R^+} \psi(x(T)) < \infty.$$

Obviously (1.1iii), (5.2) and (5.4) yield $w \in L^\infty(R^+;\, W')$, hence also (5.1i).

We call E the energy function of a given solution x of (E). If (5.3) holds, then we say that the solution has a finite energy.

Summarizing the preceding argument we get

PROPOSITION 5.1. *Suppose that* (5.2) *holds, and let x be a finite energy solution of* (E). *Then $\sigma(y) \subset Z(a)$ for every $y \in \Gamma(w)$.*

Without any further conditions on φ, ψ and f one cannot expect a solution of (E) to have a finite energy, as this is not true even in the scalar case $W = H = W' = R$. One needs some assumption connecting φ and ψ, and some global bound on f, e.g. the following:

(5.5) $$y \in \partial\varphi(x) \text{ and } z \in \partial\psi(x) \cap H \text{ implies } (y, z) \geqq 0,$$

(5.6) $$f \in L^1(R^+;\, W).$$

It remains an open question whether under the conditions (5.2), (5.5) and (5.6) every solution of (E) has a finite energy. The answer is affirmative in the special case when w in (1.1iii) satisfies $w \in L^2_{\text{loc}}(R^+;\, H)$, because then the following computation can be justified. Take the inner product in H of $w(t)$ and (1.1iv), and integrate over $[0, T]$.

This yields

$$(5.7)\quad \psi(x(T))+\int_0^T \left(w(t), \int_0^t a(t-s)w(s)\,ds\right) dt \leqq \psi(x(0))+\int_0^T (w(t), f(t))\,dt,$$

where we have used (5.5) to drop one term. By (1.1iii), (5.2) and (5.6),

$$(5.8)\qquad |(w(t), f(t))| \leqq \alpha(t)(1+\psi(x(t))) \quad \text{a.e. on } R^+,$$

where the function $\alpha(t)=K|f(t)|_W$ satisfies $\alpha \in L^1(R^+; R)$. The double integral in (5.7) can be written

$$(5.9)\qquad \int_0^T \left(w(t), \int_0^t a(t-s)w(s)\,ds\right) dt = \frac{1}{2\pi}\int_R |(\chi_{[0,T]}w)\hat{\ }(\omega)|_{\bar{H}}\, d\lambda(\omega),$$

where $\bar{H}$ is the complexification of H (the identity (5.9) is proved in [9, pp. 234–235] in the scalar case, and the same proof applies here). Recall that we without loss of generality can take $|\cdot|_H \geqq |\cdot|_W$, hence also $|\cdot|_{\bar{H}} \geqq |\cdot|_B$. Then it follows from (5.9) that

$$(5.10)\qquad Q(w, T) \leqq \int_0^T \left(w(t), \int_0^t a(t-s)w(s)\,ds\right) dt.$$

Combining (5.7), (5.8) and (5.10) we get

$$(5.11)\qquad \psi(x(T))+Q(w, T) \leqq C+\int_0^T \alpha(t)\psi(x(t))\,dt,$$

where $C=\psi(x(0))+\int_{R^+}\alpha(t)\,dt$, and, in particular,

$$\psi(x(T)) \leqq C+\int_0^T \alpha(t)\psi(x(t))\,dt.$$

Gronwall's inequality yields $\sup_{t\in R}\psi(x(t))<\infty$. Substituting this in the right hand side of (5.11) we finally obtain (5.3).

If $w \notin L^2_{\text{loc}}(R^+; H)$, then the preceding argument breaks down. However, it is still possible to verify (5.3) in some cases. In particular, the existence theorems in [1], [3]–[5] produce solutions which have a finite energy whenever a, φ, ψ and f in addition to (5.2), (5.5) and (5.6) (and a positive definite) satisfy whatever is needed to apply these theorems. This is due to the fact that one constructs a solution of (E) by solving a sequence of approximate equations (E_n), and making the solutions x_n of (E_n) converge to a solution x of (E). To each x_n corresponds a w_n, and one defines $E_n(T)$ as in (5.3), but with x, w replaced by x_n, w_n. Each w_n satisfies $w_n \in L^2_{\text{loc}}(R^+; H)$, and the approximate equations are constructed in such a way that the argument above applies, yielding

$$\sup_{n\in N, T\in R^+} E_n(T)<\infty.$$

As $n\to\infty$, it is shown that $x_n(T)\to x(T)$ weakly in W, and $w_n\to w$ weakly in $L^2(0, T; W')$ for each fixed $T\in R^+$. Thus by the weak lower semicontinuity of ψ and Q one gets (5.3), i.e. the constructed solution has a finite energy.

6. Examples. To illustrate the theory of § 5 we discuss two examples. In particular we concentrate on the verification of (5.2), which is not used in [1], [3]–[5], although it is satisfied in most of the examples given there (with the right choice of the space W). The conditions (5.2) and (5.6) combined are a special case of MacCamy's and Wong's Condition (F_0) [6].

We begin by discussing [3, Example 1]. Let Ω be a bounded open domain in R^N with smooth boundary $\partial\Omega$, and consider the integro-differential equation

$$u_t(t, x) - \Delta u(t, x) + \int_0^t a(t-s) g(u(s, x))\, ds = F(t, x) \tag{6.1}$$

for $(t, x) \in (0, \infty) \times \Omega$, together with the boundary condition

$$-\frac{\partial u}{\partial n} \in \gamma(u) \quad \text{a.e. on } (0, \infty) \times \partial\Omega \tag{6.2}$$

and the initial condition

$$u(0, x) = u_0(x), \qquad x \in \Omega \tag{6.3}$$

(here the letter x is used as a parameter, and u takes the place of the function x in (E)). The kernel a is supposed to be continuous and positive definite. It is shown in [3] that under appropriate conditions on g, F, γ and u_0 the problem (6.1)–(6.3) can be transformed into the form (E), with φ, ψ having the properties mentioned in § 1, and that (5.5) holds. Here we describe only that part of this transformation which is vital in the proof of (5.2), and we refer the reader to [3] for a full description.

We suppose that g satisfies

$$g \in C(R; R), \quad g \text{ is nondecreasing}, \quad g(0) = 0, \tag{6.4}$$

$$|g(\xi)| \leqq c_1(|\xi|^{p-1} + 1) \qquad (\xi \in R), \tag{6.5}$$

$$G(\xi) \geqq c_2|\xi|^p - c_3 \qquad (\xi \in R), \tag{6.6}$$

where c_1, c_2, c_3, p are constants, $c_1, c_2 > 0$, $c_3 \geqq 0$, $p \geqq 2$, and $G(\xi) = \int_0^\xi g(\eta)\, d\eta$ ($\xi \in R$). Choose $W = L^p(\Omega; R)$, $H = L^2(\Omega; R)$, which gives $W' = L^{p/(p-1)}(\Omega; R)$. Define

$$\psi(u) = \int_\Omega G(u(x))\, dx \qquad (u \in W). \tag{6.7}$$

Then ψ is nonnegative, convex and continuous on W, and $\partial\psi$, given by

$$\partial\psi(u) = g \circ u, \tag{6.8}$$

maps W continuously into W' (in particular, it is single-valued).

We claim that (5.2) holds with this choice of ψ, W. Take an arbitrary $u \in W$, and use (6.5)–(6.8), Minkowski's inequality and some obvious estimates to get

$$\|\partial\psi(u)\|_{W'} = \left[\int_\Omega |g(u(x))|^{p/(p-1)}\, dx\right]^{(p-1)/p}$$

$$\leqq c_1\left[\int_\Omega (|u(x)| + 1)^p\, dx\right]^{(p-1)/p} \leqq c_1[\|u\|_W + (c_4)^{1/p}]^{(p-1)},$$

$$\psi(u) = \int_\Omega G(u(x))\, dx \geqq \int_\Omega (c_2|u(x)|^p - c_3)\, dx = c_2\|u\|_W^p - c_3 c_4,$$

where $c_4 = \int_\Omega dx$. This implies

$$\limsup_{\|u\|_W \to \infty} \|\partial\psi(u)\|_{W'}/\psi(u) < \infty, \tag{6.9}$$

which together with the nonnegativity of ψ and the fact that $\partial\psi$ is bounded on bounded sets yields (5.2).

Applying Proposition 5.1 to this particular example we conclude that every finite energy solution of (6.1)–(6.3) (i.e. a solution for which (1.1), (5.3) hold) satisfies $\sigma(y)\subset Z(a)$ $(y\in\Gamma(g\circ u))$, where we consider $g\circ u$ as a map from R^+ into $L^{p/(p-1)}(\Omega; R)$. In particular, if $Z(a)=\varnothing$ (which is true e.g. when $a(t)=e^{-t}(t\in R^+)$), then the left-translates $g(u(t+s,x))$ $(t,s\in R^+; x\in\Omega)$ of $g\circ u$ tend weak* to zero in $L^2_\infty(R; L^{p/(p-1)}(\Omega; R))$ as $t\to\infty$. Moreover, if in addition $F\in L^1(R^+; L^p(\Omega; R))$, then any solution constructed by an approximating method of the type outlined in § 5 has a finite energy.

Remark 6.1. The bound (6.6) is not needed in the proof of existence of solutions of (6.1)–(6.3) (see [3]). On the other hand, the existence proof in [3] instead uses a number of additional assumptions not mentioned above.

As a second example we consider the equation (cf. [4])

$$u_t(t,x)-\int_0^t a(t-s)g(u_x(s,x))_x\,ds=F(t,x) \tag{6.10}$$

for $t\in R^+$, $0\leqq x\leqq 1$, with the boundary and initial conditions

$$u(t,0)=u(t,1)=0\quad (t\in R^+),\qquad u(0,x)=u_0(x)\quad (0\leqq x\leqq 1). \tag{6.11}$$

Again a is continuous and positive definite. We suppose that g satisfies (6.4)–(6.6) with $p=2$. This time take $W=H_0^1(0,1;R)$, $H=L^2(0,1;R)$, which gives $W'=H^{-1}(0,1;R)$. Use the norm $\|u\|_W=[\int_0^1|u_x|^2\,dx]^{1/2}$ in W. Define

$$\psi(u)=\int_0^1 G(u_x)\,dx\qquad (u\in W). \tag{6.12}$$

Then ψ is nonnegative, convex and continuous on W, and it has a lower semicontinuous extension to H (define $\psi(u)=\infty$ $(u\in H\setminus W)$). The subgradient $\partial\psi\colon W\to W'$, given by

$$\langle\partial\psi(u),v\rangle=\int_0^1 g(u_x)v_x\,dx\qquad (u,v\in W) \tag{6.13}$$

is single-valued and defined everywhere on W. Another way to write (6.13) which corresponds more closely to (6.10) is

$$\partial\psi(u)=-(g(u_x))_x\qquad (u\in W), \tag{6.14}$$

where the differentiations should be interpreted in the distribution sense. It follows from (6.5)–(6.6) (with $p=2$), (6.12), (6.13), the Schwarz inequality and Minkowski's inequality that

$$\|\partial\psi(u)\|_{W'}=\sup_{\|v\|_W=1}\langle\partial\psi(u),v\rangle=\sup_{\|v\|_W=1}\int_0^1 g(u_x)v_x\,dx$$
$$\leqq\left[\int_0^1 (g(u_x))^2\,dx\right]^{1/2}\leqq c_1(\|u\|_W+1),$$
$$\psi(u)=\int_0^1 G(u_x)\,dx\geqq c_2\|u\|_W^2-c_3.$$

As in the previous example we conclude that (5.2) holds.

Applying Proposition 5.1 we find that every finite energy solution of (6.10)–(6.11) satisfies $\sigma(y)\subset Z(a)$ $(y\in\Gamma((g(u_x))_x))$. If in addition $F\in L^1(R^+; H_0^1(0,1;R))$, then e.g. the solution constructed in [4] is a finite energy solution.

Remark 6.2. Some additional assumptions are needed before one can apply the existence result in [4]. In particular, the derivative of the kernel a is not allowed to be bounded, so e.g. the case $a(t)\equiv 1$ (the nonlinear wave equation) is excluded.

REFERENCES

[1] V. BARBU, *Integro-differential equations in Hilbert spaces*, An. Şti. Univ. "Al. I. Cuza" Iaşi Secţ. I a Mat., 19 (1973), pp. 365–383.

[2] H. BRÉZIS, *Operateurs Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert*, North-Holland, Amsterdam, 1973.

[3] M. G. CRANDALL, S.-O. LONDEN AND J. A. NOHEL, *An abstract nonlinear Volterra integrodifferential equation*, J. Math. Anal. Appl., to appear.

[4] S.-O. LONDEN, *An existence result on a Volterra equation in a Banach space*, Trans. Amer. Math. Soc., to appear.

[5] ———, *On a Volterra integrodifferential equation in a Banach space*, Tech. Summary Rep. 1558, MRC, Univ. of Wisconsin, Madison, 1975.

[6] R. C. MACCAMY AND J. S. W. WONG, *Stability theorems for some functional equations*, Trans. Amer. Math. Soc., 164 (1972), pp. 1–37.

[7] W. RUDIN, *Functional Analysis*, McGraw-Hill, New York, 1973.

[8] L. SCHWARTZ, *Théorie des distributions a valeurs vectorielles*, Ann. Inst. Fourier (Grenoble), 7 (1957), pp. 1–141; 8 (1958), pp. 1–209.

[9] O. J. STAFFANS, *Positive definite measures with applications to a Volterra equation*, Trans. Amer. Math. Soc., 218 (1976), pp. 219–237.

[10] ———, *Tauberian theorems for a positive definite form, with applications to a Volterra equation*, Ibid., 218 (1976), pp. 239–259.

[11] ———, *An asymptotic problem for a positive definite operator-valued Volterra kernel*, this Journal, 9 (1978), pp. 855–866.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0012 $01.00/0

ESTIMATES ON THE EXISTENCE REGIONS OF PERTURBED PERIODIC SOLUTIONS*

M. FARKAS†

Abstract. The n dimensional perturbed system of differential equations $\dot{x}=f(x)+\mu g(t/\tau, x)$ is considered. It is assumed that g is periodic in the variable t with period τ and that the unperturbed system $\dot{x}=f(x)$ has a nonconstant isolated periodic solution with period τ_0. Explicit bounds of a region are given in which the "small parameter" μ may vary. To each μ in this region there belongs a period τ such that if this value of τ is written into the perturbed system it has a periodic solution with this period. In the estimates data gained from the right hand side of the system and from solutions of a variational system of the unperturbed system only are used and no information about the solutions of the perturbed system is needed.

1. Introduction. In papers [1] and [2] the author has considered controllably periodic perturbations of an autonomous system. The perturbed system was assumed in the form

$$\dot{x}=f(x)+\mu g\left(\frac{t}{\tau}, x, \mu, \tau\right) \tag{1.1}$$

where $x \in R^n$, the dot means differentiation with respect to $t \in R$, μ is a "small parameter" and τ is a real parameter. It was assumed that the unperturbed system $\dot{x}=f(x)$ has a nonconstant, isolated periodic solution p with period τ_0 and that g is periodic in the variable t with period τ. An auxiliary variable, the "initial phase" φ of the solutions has been introduced and under fairly general conditions it has been proved that to each "small enough" value of $|\mu|$ and $|\varphi|$ there belongs a unique initial value and a period τ such that if this value of τ is written into (1.1) the system has a periodic solution with this period. The main condition is that the number 1 is a simple characteristic multiplier of the variational system corresponding to the solution p of the unperturbed system. This condition is sufficient for p to have an isolated path (see [3]).

The aim of the present paper is to establish explicit bounds for the parameters among which the existence of the perturbed periodic solutions is ensured. Similar investigations have been carried out in cases different from ours by D. C. Lewis [7], H. I. Freedman [4], [5] and others (see the references in papers [4] and [5]). Our method is related to the method applied by Freedman; namely, an implicit function theorem is used. However, in spite of the fact that a "critical case" is treated here, the use of vanishing Jacobians has been avoided. This has been achieved at the expense of establishing the existence of functions of two real variables instead of a single one. As a consequence a more complex implicit function theorem had to be applied. At the same time we had to establish the region where the Jacobian does not vanish, a condition which in paper [4] was taken for granted.

In § 2 for better understanding we collect (in a somewhat modified form) the main results of the references that are to be used in the subsequent sections. The tough job of giving estimates for the differences of the solutions and of the derivatives of the solutions is accomplished in § 3. The main results are contained in § 4.

Throughout the paper if x is an element of R^n, its coordinates are denoted by the same letter with subscripts, i.e. $x=(x_1, x_2, \cdots, x_n)$. At the same time in matrix

* Received by the editors December 28, 1976.
† Department of Mathematics, Budapest University of Technology, 1521 Budapest, Hungary.

algebraic operations the elements of R^n are dealt with as column vectors. For the norm of a vector $u = \text{col}(u_1, u_2, \cdots, u_m)$ we shall use $|u| = \max_i |u_i|$. If D is an n by m matrix, i.e. $D = [d_{ik}]$ $(i = 1, 2, \cdots, n;\ k = 1, 2, \cdots, m)$, its norm is defined by $|D| = m \max_{i,k} |d_{ik}|$. It is easy to see that besides some other basic properties $|Du| \leqq |D||u|$. If in particular the matrices A and B are quadratic matrices of order n $(m = n)$, then $|AB| \leqq |A||B|$. Naturally, some other norms could be used as well and, in fact, in some cases might yield sharper results. We shall have to apply "tensors with three subscripts". Thus, if $E = [e_{ikl}]$, $(i, k, l = 1, 2, \cdots, n)$ its norm is defined by $|E| = n^2 \max_{i,k,l} |e_{ikl}|$. The inner product of E and the vector $x \in R^n$ is the matrix

$$Ex = \left[\sum_{l=1}^{n} e_{ikl}x_l\right] \qquad (i, k = 1, 2, \cdots, n).$$

If $y \in R^n$ then

$$Exy = \left(\sum_{k,l=1}^{n} e_{1kl}x_ly_k, \cdots, \sum_{k,l=1}^{n} e_{nkl}x_ly_k\right) \in R^n.$$

It is easy to see that $|Ex| \leqq |E||x|$ and $|Exy| \leqq |E||x||y|$. Finally, if the function $f\colon R^n \to R^n$ belongs to the C^2 class we shall apply the following notations:

$$f'_x(x) = [f'_{ix_k}(x)] \qquad (i, k = 1, 2, \cdots, n)$$

and

$$f''_{xx}(x) = [f''_{ix_kx_l}(x)] \qquad (i, k, l = 1, 2, \cdots, n),$$

the former being the "Jacobi matrix", the latter a tensor with three subscripts whose elements are the second partial derivatives of the coordinates of f.

2. Some results to be applied. The following implicit function theorem, due to D. C. Lewis [6], is presented with a new proof here.

THEOREM 2.A. *Consider the system of n equations with $n + m$ unknowns*

$$z(u, v) = 0 \tag{2.1}$$

where $u = (u_1, \cdots, u_m)$, $v = (v_1, \cdots, v_n)$ and with the notation

$$W = \{(u, v) \in R^m \times R^n\colon |u| \leqq a, |v| \leqq \beta\}$$

with some constants $a > 0$, $\beta > 0$, the function $z\colon W \to R^n$ belongs to the C^2_W class; assume that $z(0, 0) = 0$ and $\det z'_v(u, v) \neq 0$, $(u, v) \in W$; *the linear system of equations*

$$z'_v(u, v)d^h + z'_{u_h}(u, v) = 0, \tag{2.2}$$

clearly, defines the functions $d^h\colon W \to R^n$ for each $h = 1, 2, \cdots, m$ uniquely and $d^h =$ $\text{col}(d^h_1, d^h_2, \cdots, d^h_n) \in C^1_W$; *if we construct the n by m matrix $D = [d^1, d^2, \cdots, d^m]$ out of the column vectors d^h there exists a $\Delta > 0$ such that*

$$|D(u, v)| \leqq \Delta, \qquad (u, v) \in W; \tag{2.3}$$

let $c = (c_1, \cdots, c_m)$ be any vector of positive coordinates $(c_h > 0,\ h = 1, 2, \cdots, m)$ such that

$$|c| = \max_h |c_h| < \min\left(a, \frac{\beta}{\Delta}\right) \tag{2.4}$$

and $U=\{u\in R^m: |u|<|c|\}$; *then there exists one and only one function* $w: U\to R^n$ *such that* $w\in C_U^1$, $|w(u)|<\beta$ if $u\in U$,

$$w(0)=0 \tag{2.5}$$

and

$$z(u, w(u))\equiv 0, \qquad u\in U. \tag{2.6}$$

Proof. If a function $w: U\to R^n$ is in the C_U^1 class, $|w(u)|<\beta$ for $u\in U$, and w satisfies the condition (2.6), then differentiating the latter identity with respect to u_h, we get

$$z'_v(u, w(u))w'_{u_h}(u)+z'_{u_h}(u, w(u))\equiv 0 \qquad (h=1, 2, \cdots, m),$$

i.e. $w'_{u_h}(u)\equiv d^h(u, w(u))$. Conversely, if a function w is in the C_U^1 class, $|w(u)|<\beta$ in U, $w(0)=0$, and w satisfies the system of partial differential equations

$$v'_{u_h}=d^h(u, v) \qquad (h=1, 2, \cdots, m), \tag{2.7}$$

then, clearly, it satisfies (2.6) as well.

For system (2.7) the "condition of complete integrability"

$$d^{h\prime}_{u_l}+d^{h\prime}_v d^l-d^{l\prime}_{u_h}-d^{l\prime}_v d^h=0 \qquad (h, l=1, 2, \cdots, m) \tag{2.8}$$

holds in W. To prove (2.8) assume that the solution of (2.2) has been substituted into (2.2) and differentiate this identity first with respect to u_l, then with respect to v. In the latter case we multiply the derived identity by the column vector d^l from the right. In the first case we get:

$$z''_{vu_l}d^h+z'_v d^{h\prime}_{u_l}+z''_{u_hu_l}=0;$$

in the second:

$$z''_{vv}d^hd^l+z'_v d^{h\prime}_v d^l+z''_{u_hv}d^l=0.$$

(Here, as it was pointed out in the Introduction,

$$z''_{vv}d^hd^l=\sum_{i,j=1}^{n} z''_{v_iv_j}d^h_j d^l_i,$$

which is considered as a column vector. The writing out of the arguments has been suppressed everywhere.) Adding the last two identities, we obtain

$$z''_{vu_l}d^h+z'_v d^{h\prime}_{u_l}+z''_{u_hu_l}+z''_{vv}d^hd^l+z'_v d^{h\prime}_v d^l+z''_{u_hv}d^l=0.$$

Interchanging the roles of h and l and subtracting the two identities yields

$$z'_v(d^{h\prime}_{u_l}+d^{h\prime}_v d^l-d^{l\prime}_{u_h}-d^{l\prime}_v d^h)=0.$$

Since the matrix z'_v is regular in W, equation (2.8) follows.

In what follows we are going to show that (2.7) has a solution $w\in C_U^l$ for which the conditions $|w(u)|<\beta$ in U and $w(0)=0$ hold.

Let u^* be an arbitrary point in U and consider the system of ordinary differential equations

$$\frac{dx}{dt}=D(u^*t, x)u^* \tag{2.9}$$

with the initial condition $x(0, u^*)=0$. The right hand side of (2.9) where the vector u^* occurs as a parameter is continuously differentiable for $u^*\in U$ (i.e. $|u^*|<|c|$), $|x|<\beta$

and $|t| < a/|c|$ (it is to be recalled that by (2.4) $a/|c| > 1$). Let us denote the supremum of the right hand side of (2.9) by M. According to (2.4)

$$M = \sup_{\substack{|u^*|<|c| \\ |t|<a/|c| \\ |x|<\beta}} |D(u^*t, x)u^*| \leqq \sup_{\substack{|u^*|<|c| \\ |t|<a/|c| \\ |x|<\beta}} |D(u^*t, x)|\,|u^*| \leqq \Delta|c| < \beta.$$

By the local existence and uniqueness theorem system (2.9) has a unique solution $x(t, u^*)$ satisfying the initial condition $x(0, u^*) = 0$ defined in the interval $(-\alpha, \alpha)$ where

$$\alpha = \min\left(\frac{a}{|c|}, \frac{\beta}{M}\right) > 1$$

and $|x(t, u^*)| < \beta$ for $t \in (-\alpha, \alpha)$. As it is well known this solution is a continuously differentiable function of the parameter u^* for all $t \in (-\alpha, \alpha)$ and $u^* \in U$.

Now we are in the position to construct the required solution of (2.7). Let us define a function $w\colon U \to R^n$ the following way:

$$w(u^*) = x(1, u^*), \qquad u^* \in U. \tag{2.10}$$

Clearly $w \in C^1_U$, $|w(u^*)| < \beta$ in U and $w(0) = x(1, 0) = 0$. We have to show that the function defined by (2.10) is a solution of the system (2.7).

For this purpose let us differentiate the identity

$$\dot{x}(t, u^*) \equiv D(u^*t, x(t, u^*))u^*$$

with respect to the coordinates u^*_h $(h = 1, 2, \cdots, m)$ of the vector parameter u^*:

$$\dot{x}'_{u^*_h}(t, u^*) \equiv \left(D'_{u_h}(u^*t, x(t, u^*))t + \sum_{i=1}^{n} D'_{v_i}(u^*t, x(t, u^*))x'_{iu^*_h}(t, u^*)\right)u^*$$
$$+ d^h(u^*t, x(t, u^*)).$$

If we take into consideration also that $x(0, u^*) \equiv 0$ and hence $x'_{u^*_h}(0, u^*) = 0$ we see that the function $x'_{u^*_h}(t, u^*)$ is a solution of the inhomogeneous linear system

$$\begin{aligned} \dot{y} &= \sum_{i=1}^{n} D'_{v_i}(u^*t, x(t, u^*))u^* y_i \\ &\quad + D'_{u_h}(u^*t, x(t, u^*))u^*t + d^h(u^*t, x(t, u^*)) \end{aligned} \tag{2.11}$$

satisfying also the initial condition $y(0) = 0$. At the same time by substituting the function $y = td^h(u^*t, x(t, u^*))$ into (2.11) and applying (2.9) and (2.8) we can see that the latter function satisfies the same system and, clearly, the same initial condition. Thus

$$x'_{u^*_h}(t, u^*) \equiv td^h(u^*t, x(t, u^*)).$$

Applying the last identity at $t = 1$, we get

$$w'_{u^*_h}(u^*) = x'_{u^*_h}(1, u^*) = d^h(u^*, x(1, u^*)) = d^h(u^*, w(u^*)),$$
$$(h = 1, 2, \cdots, m).$$

With the star in the last identity dropped, this shows that the function defined in (2.10) (which is to be read also without the star) satisfies the system (2.7) of partial differential equations. Thus the existence of a function w satisfying all the requirements of the theorem in the whole region U has been proved.

The uniqueness of the function $w: U \to R^n$ can easily be proved by some standard method. This part of the proof is presented here for sake of completeness. It is, clearly, sufficient to prove that (2.7) has only one solution satisfying the initial condition $v(0)=0$ and $|v(u)|<\beta$ in U. Assume that $w: U \to R^n$ is such a solution and define a function of a single variable by

$$w^1(u_1)=w(u_1, 0, \cdots, 0), \qquad |u_1|<|c|.$$

This function, obviously satisfies the system of ordinary differential equations

$$\frac{dw^1(u_1)}{du_1}=d^1(u_1, 0, \cdots, 0, w^1(u_1)), \qquad |u_1|<|c|,$$

and $w^1(0)=0$. Thus, w^1 is unique and hence the values of w are also uniquely determined for $|u_1|<|c|$, $u_2=\cdots=u_m=0$. Let $|u_1^*|<|c|$ and consider the function of a single variable

$$w^2(u_2)=w(u_1^*, u_2, 0, \cdots, 0), \qquad |u_2|<|c|.$$

This function, obviously, satisfies the system

$$\frac{dw^2(u_2)}{du_2}=d^2(u_1^*, u_2, 0, \cdots, 0, w^2(u_2)), \qquad |u_2|<|c|$$

and the initial condition $w^2(0)=w^1(u_1^*)$. Thus w^2 is unique and hence the values of w are also uniquely determined for $|u_1|<|c|$, $|u_2|<|c|$, $u_3=\cdots=u_m=0$. If we proceed further in a similar manner, the uniqueness of the solution in U follows, and this completes the proof.

In the rest of this section a theorem due to A. Ostrowski is quoted. Its proof can be found in [8]. Here the theorem is presented in a slightly modified form but these minor changes effect the proof only slightly. The theorem is concerned with the question, "How much does the solution of a system of linear (algebraic) equations vary if the coefficients and the constant terms are perturbed?" At the same time it gives bounds among which the elements of a regular matrix can be varied without becoming singular.

Consider the linear system

$$Ax=b \tag{2.12}$$

where $A=[a_{ik}]$ is a regular n by n quadratic matrix, $b=\text{col}(b_1, \cdots, b_n)$, a column vector of dimension n, and $x=\text{col}(x_1, \cdots, x_n)$, the column vector of solutions. A perturbation of the matrix A and the vector b is denoted by $\delta A=[\delta a_{ik}]$ and $\delta b=\text{col}(\delta b_1, \cdots, \delta b_n)$, respectively; i.e. the system

$$(A+\delta A)(x+\delta x)=b+\delta b \tag{2.13}$$

is considered whose solution is denoted by $x+\delta x=\text{col}(x_1+\delta x_1, \cdots, x_n+\delta x_n)$. System (2.13) is multiplied by A^{-1} from the left; this yields

$$(I+A^{-1}\,\delta A)(x+\delta x)=A^{-1}b+A^{-1}\,\delta b$$

where I is the n by n unit matrix.

THEOREM 2.B. *If* $|A^{-1}\delta A|<1$, *then the matrix* $A+\delta A$ *is regular and we have for the difference* δx *of the solutions of* (2.13) *and* (2.12)

$$|\delta x| \leqq |A^{-1}|\left(|\delta b|+(|b|+|\delta b|)\frac{|A^{-1}||\delta A|}{1-|A^{-1}\,\delta A|}\right). \tag{2.14}$$

The only assumption, $|A^{-1}\,\delta A|<1$, under which the inequality (2.14) is proved by Ostrowski is obviously valid if the perturbation $|\delta A|$ is small enough, more precisely if

$$|\delta A|<\frac{1}{|A^{-1}|}. \tag{2.15}$$

This inequality is also a sufficient condition for the regularity of the matrix $A+\delta A$.

3. Estimates for the variations of perturbed solutions and their derivatives. Let Ω be an open and bounded ball in R^n with center in the origin, assume that the functions $f\colon \bar\Omega\to R^n$ and $g\colon R\times\bar\Omega\to R^n$ belong to the C^2 class in the closure $\bar\Omega$ of the ball Ω, and consider the perturbed system of differential equations

$$\dot x=f(x)+\mu g\left(\frac{t}{\tau},x\right) \tag{3.1}$$

where $\mu\in R$, $\tau>0$, and it is also assumed that g is periodic in the variable t with period τ, i.e.

$$g(s+1,x)\equiv g(s,x),\qquad s\in R,\quad x\in\bar\Omega.$$

The function g is assumed in a somewhat simpler form than it was originally in [1]. Keeping to the original more general form of g would not cause any theoretical difficulties; however, it would increase the length of the calculations and formulae.

Along with (3.1) consider the unperturbed system

$$\dot x=f(x). \tag{3.2}$$

It will be assumed that (3.2) has a nonconstant periodic solution $p\colon R\to\Omega$ with period $\tau_0>0$ and, further, that the number 1 is a simple characteristic multiplier of the variational system

$$\dot y=f'_x(p(t))y. \tag{3.3}$$

The notations $p^0=p(0)=(p_1^0,\cdots,p_n^0)$ and $\dot p^0=\dot p(0)=\mathrm{col}\,(\dot p_1^0,\cdots,\dot p_n^0)$ are introduced. It will be assumed, without loss of generality, that

$$\dot p_1^0\neq 0,\qquad \dot p_i^0=0,\qquad (i=2,3,\cdots,n). \tag{3.4}$$

Vectors orthogonal to $\dot p^0$ will be denoted by h. Because of (3.4) the first coordinate of such a vector is zero: $h=(0,h_2,\cdots,h_n)$. The solution of (3.1) that assumes the value p^0+h (in the hyperplane passing through p^0 and orthogonal to $\dot p^0$) at $t=\varphi$ will be denoted by $x(t;\varphi,p^0+h,\mu,\tau)$. As has been pointed out in [1], for small enough $|\mu|$ all solutions with initial values close enough to $t=0$, $x=p^0$ can be written in this form. The condition of periodicity (with period τ) is

$$z(\mu,\varphi,\tau,h)\overset{\text{def}}{=}x(\varphi+\tau;\varphi,p^0+h,\mu,\tau)-p^0-h=0. \tag{3.5}$$

The Jacobi matrix of the function z with respect to the variables $\tau,h_2,\cdots,h_n$ is

$$\begin{aligned}J(\mu,\varphi,\tau,h)=[&\dot x(\varphi+\tau;\varphi,p^0+h,\mu,\tau)+x'_\tau(\varphi+\tau;\varphi,p^0+h,\mu,\tau),\\&x'_{h_2}(\varphi+\tau;\varphi,p^0+h,\mu,\tau)-e^2,\cdots,x'_{h_n}(\varphi+\tau;\varphi,p^0+h,\mu,\tau)-e^n]\end{aligned}$$

where $e^2 = \text{col}\,(0, 1, 0, \cdots, 0), \cdots, e^n = \text{col}\,(0, 0, \cdots, 0, 1)$. It has been proved in [1] that

$$\det J(0, 0, \tau_0, 0) \neq 0, \tag{3.6}$$

assuring the existence of uniquely determined continuously differentiable functions $\tau(\mu, \varphi)$ and $h(\mu, \varphi)$ such that the function $x(t; \varphi, p^0 + h(\mu, \varphi), \mu, \tau(\mu, \varphi))$ is a periodic solution with period $\tau(\mu, \varphi)$ of system

$$\dot{x} = f(x) + \mu g\left(\frac{t}{\tau(\mu, \varphi)}, x\right),$$

and $\tau(0, 0) = \tau_0$, $h(0, 0) = 0$. The aim of the present paper is to give an estimate for the region in which the variables μ and φ may vary. For this purpose estimates are needed for the norm of the difference $J(\mu, \varphi, \tau, h) - J(0, 0, \tau_0, 0)$ and for Δ occurring in (2.3).

The following quantities are considered to be known:

$$\begin{aligned} &F_0 = \max_{x \in \bar{\Omega}} |f(x)|, && G_0 = \max_{\substack{x \in \bar{\Omega} \\ s \in R}} |g(s, x)|, \\ &F_1 = \max_{x \in \bar{\Omega}} |f'_x(x)|, && G_1 = \max_{\substack{x \in \bar{\Omega} \\ s \in R}} |g'_x(s, x)|, \\ &F_2 = \max_{x \in \bar{\Omega}} |f''_{xx}(x)|, && G_s = \max_{\substack{x \in \bar{\Omega} \\ s \in R}} |g'_s(s, x)|. \end{aligned} \tag{3.7}$$

$Y(t)$ will denote the fundamental matrix solution of the unperturbed variational system (3.3) that assumes the unit matrix at $t = 0$, i.e. $I = Y(0)$. This matrix is considered to be known along with its inverse $Y^{-1}(t)$ and

$$K = \max_{t \in [-\tau_0/2, \tau_0]} |Y(t)|, \qquad K_{-1} = \max_{t \in [-\tau_0/2, \tau_0]} |Y^{-1}(t)|. \tag{3.8}$$

It is to be noted that the first column of $Y(t)$ is $(1/\dot{p}_1^0)\dot{p}(t)$ and the ith one ($i = 2, 3, \cdots, n$) is $x'_{h_i}(t; 0, p^0, 0, \tau_0)$. Thus

$$P = \max |\dot{p}(t)| \leq \frac{K}{n} |\dot{p}_1^0|, \ |x'_{h_i}(t; 0, p^0, 0, \tau_0)| \leq \frac{K}{n}, \qquad t \in [-\tau_0/2, \tau_0], \quad (i = 2, 3, \cdots, n). \tag{3.9}$$

In the following estimates, of course, each F, G and K can be replaced by an upper bound.

The path of the periodic solution p will be denoted by $\gamma = \{x \in R^n : x = p(t), t \in [0, \tau_0]\}$ and its distance from the boundary of the ball Ω by

$$\sigma = \text{dist}\,(\gamma, \text{front}\,\Omega) > 0.$$

The domain in which the "initial phase" φ and the period τ varies will be restricted from now on by

$$|\varphi| < \tau_0/2, \qquad |\tau - \tau_0| < \tau_0/2. \tag{3.10}$$

Since τ_0 need not be the least positive period of the function p, conditions (3.10) do not restrict generality, at least, as far as periods greater than τ_0 are concerned.

First of all we need bounds for $|\mu|$ and $|h|$ such that for arbitrary φ and τ satisfying (3.10) the solution $x(t;\varphi,p^0+h,\mu,\tau)$ of (3.1) should be defined and its path contained in Ω for $t\in[\varphi,\varphi+\tau]$.

LEMMA 3.1. *If μ and h are such that*

$$\tfrac{3}{2}G_0\tau_0|\mu|+|h|<\sigma\exp\left(-\tfrac{3}{2}F_1\tau_0\right), \tag{3.11}$$

then the solution $x(t;\varphi,p^0+h,\mu,\tau)$ is defined and $x(t;\varphi,p^0+h,\mu,\tau)\in\Omega$ for all φ and τ satisfying (3.10) *and $t\in[\varphi,\varphi+\tau]$.*

Proof. Consider the identities

$$\dot x(t;\varphi,p^0+h,\mu,\tau)\equiv f(x(t;\varphi,p^0+h,\mu,\tau))+\mu g\left(\frac{t}{\tau},x(t;\varphi,p^0+h,\mu,\tau)\right), \tag{3.12}$$

$$\dot p(t-\varphi)\equiv f(p(t-\varphi)). \tag{3.13}$$

Integrate the difference of these identities from φ to $t>\varphi$:

$$\begin{aligned}x(t;\varphi,p^0+h,\mu,\tau)-p^0-h-p(t-\varphi)+p^0\\ \equiv\int_\varphi^t\Big(&f(x(u;\varphi,p^0+h,\mu,\tau))-f(p(u-\varphi))\\ &+\mu g\left(\frac{u}{\tau},x(u;\varphi,p^0+h,\mu,\tau)\right)\Big)\,du.\end{aligned}$$

Provided that the path of the solution x is in Ω in the interval $[\varphi,t)$ we get from here

$$\begin{aligned}&|x(t;\varphi,p^0+h,\mu,\tau)-p(t-\varphi)|\\ &\leqq|h|+\int_\varphi^t(F_1|x(u;\varphi,p^0+h,\mu,\tau)-p(u-\varphi)|+G_0|\mu|)\,du.\end{aligned}$$

Applying Gronwall's lemma, we obtain

$$|x(t;\varphi,p^0+h,\mu,\tau)-p(t-\varphi)|\leqq(G_0|\mu|(t-\varphi)+|h|)\exp(F_1(t-\varphi)).$$

The solution $x(t;\varphi,p^0+h,\mu,\tau)$ is defined and its path is contained in Ω as long as

$$\operatorname{dist}(x(t;\varphi,p^0+h,\mu,\tau),\gamma)<\sigma. \tag{3.14}$$

The left hand side of the last inequality is less than the left hand side of the previous one. Since $0\leqq t-\varphi\leqq\tau<\frac{3}{2}\tau_0$, inequality (3.14) follows from (3.11) for all $t\in[\varphi,\varphi+\tau]$, and this completes the proof.

From now on it will be assumed that μ and h satisfy (3.11). The following three lemmata yield estimates for the variations of the columns of matrix $J(\mu,\varphi,\tau,h)$. The abbreviated notations

$$\delta x'_{h_i}(t)=x'_{h_i}(t;\varphi,p^0+h,\mu,\tau)-x'_{h_i}(t+\tau_0-\tau-\varphi;0,p^0,0,\tau_0),\qquad(i=2,3,\cdots,n),$$

$$\delta f'_x(t)=f'_x(x(t;\varphi,p^0+h,\mu,\tau))-f'_x(p(t+\tau_0-\tau-\varphi))$$

will be used. Obviously, $\delta x'_{h_i}(\varphi+\tau)$ is the difference between the ith columns of the matrices $J(\mu,\varphi,\tau,h)$ and $J(0,0,\tau_0,0)$.

LEMMA 3.2. *If conditions* (3.10) *and* (3.11) *are satisfied then*

$$|\delta x'_{h_i}(\varphi+\tau)| \leqq (\tfrac{3}{2}(\tfrac{3}{2}G_0\tau_0|\mu|+|h|+F_0|\tau-\tau_0|)F_2\tau_0 \exp[\tfrac{3}{2}F_1\tau_0] + \tfrac{3}{2}G_1\tau_0|\mu|+F_1|\tau-\tau_0|)\frac{K}{n}\exp[\tfrac{3}{2}\tau_0(F_1+G_1|\mu|)], \quad (i=2,3,\cdots,n). \tag{3.15}$$

Proof. Differentiating the identity (3.12) with respect to h_i we get

$$\dot{x}'_{h_i}(t;\varphi,p^0+h,\mu,\tau) \equiv \Big(f'_x(x(t;\varphi,p^0+h,\mu,\tau)) + \mu g'_x\Big(\frac{t}{\tau}, x(t;\varphi,p^0+h,\mu,\tau)\Big)\Big)x'_{h_i}(t;\varphi,p^0+h,\mu,\tau). \tag{3.16}$$

Substituting $\varphi=0$, $h=0$, $\mu=0$, $\tau=\tau_0$ and writing $t+\tau_0-\tau-\varphi$ for t, we obtain the identity

$$\dot{x}'_{h_i}(t+\tau_0-\tau-\varphi;0,p^0,0,\tau_0) \equiv f'_x(p(t+\tau_0-\tau-\varphi))x'_{h_i}(t+\tau_0-\tau-\varphi;0,p^0,0,\tau_0).$$

Subtracting the last identity from the previous one and integrating the difference from φ to $t>\varphi$, we get

$$\begin{aligned} x'_{h_i}(t;\varphi,p^0+h,\mu,\tau) &- x'_{h_i}(t+\tau_0-\tau-\varphi;0,p^0,0,\tau_0) \\ \equiv\ & x'_{h_i}(\varphi;\varphi,p^0+h,\mu,\tau) - x'_{h_i}(\tau_0-\tau;0,p^0,0,\tau_0) \\ &+ \int_\varphi^t \Big(\Big(f'_x(x(u;\varphi,p^0+h,\mu,\tau)) \\ &\quad + \mu g'_x\Big(\frac{u}{\tau}, x(u;\varphi,p^0+h,\mu,\tau)\Big)\Big)\Big(x'_{h_i}(u;\varphi,p^0+h,\mu,\tau) \\ &\quad - x'_{h_i}(u+\tau_0-\tau-\varphi;0,p^0,0,\tau_0)\Big) \\ &\quad + \Big(f'_x(x(u;\varphi,p^0+h,\mu,\tau)) + \mu g'_x\Big(\frac{u}{\tau}, x(u;\varphi,p^0+h,\mu,\tau)\Big) \\ &\quad - f'_x(p(u+\tau_0-\tau-\varphi))\Big)x'_{h_i}(u+\tau_0-\tau-\varphi;0,p^0,0,\tau_0)\Big)\,du, \end{aligned}$$

or applying the notations previously introduced, we get

$$\begin{aligned} \delta x'_{h_i}(t) \equiv \delta x'_{h_i}(\varphi) + \int_\varphi^t \Big(&\Big(f'_x(x(u;\varphi,p^0+h,\mu,\tau)) \\ &+ \mu g'_x\Big(\frac{u}{\tau}, x(u;\varphi,p^0+h,\mu,\tau)\Big)\Big)\,\delta x'_{h_i}(u) \\ &+ \Big(\delta f'_x(u) + \mu g'_x\Big(\frac{u}{\tau}, x(u;\varphi,p^0+h,\mu,\tau)\Big)\Big) \\ &\cdot x'_{h_i}(u+\tau_0-\tau-\varphi;0,p^0,0,\tau_0)\Big)\,du. \end{aligned} \tag{3.17}$$

In order to continue, an estimate is needed for $|\delta x'_{h_i}(\varphi)|$ and $|\delta f'_x(t)|$. Clearly

$$\begin{aligned} \delta x'_{h_i}(\varphi) &= x'_{h_i}(\varphi;\varphi,p^0+h,\mu,\tau) - x'_{h_i}(\tau_0-\tau;0,p^0,0,\tau_0) \\ &= e^i - x'_{h_i}(\tau_0-\tau;0,p^0,0,\tau_0) \\ &= x'_{h_i}(0;0,p^0,0,\tau_0) - x'_{h_i}(\tau_0-\tau;0,p^0,0,\tau_0). \end{aligned}$$

As is well known, $x'_{h_i}(t; 0, p^0, 0, \tau_0)$ satisfies system (3.3), thus

$$|\delta x'_{h_i}(\varphi)| = \left| \int_0^{\tau_0-\tau} f'_x(p(t)) x'_{h_i}(t; 0, p^0, 0, \tau_0)\, dt \right|.$$

Hence applying (3.7) and (3.9), we have

$$|\delta x'_{h_i}(\varphi)| \leqq F_1 \frac{K}{n} |\tau - \tau_0|. \tag{3.18}$$

Applying mean value theorems for the elements of the matrix $\delta f'_x(t)$ we easily obtain

$$|\delta f'_x(t)| \leqq F_2 |x(t; \varphi, p^0 + h, \mu, \tau) - p(t + \tau_0 - \tau - \varphi)|. \tag{3.19}$$

To estimate the right hand side, one may proceed as in the proof of Lemma 3.1. Besides (3.12) one has to consider, instead of (3.13), the identity

$$\dot p(t + \tau_0 - \tau - \varphi) \equiv f(p(t + \tau_0 - \tau - \varphi)).$$

The procedure yields

$$\begin{aligned} &|x(t; \varphi, p^0 + h, \mu, \tau) - p(t + \tau_0 - \tau - \varphi)| \\ &\quad \leqq (G_0 |\mu| (t - \varphi) + |h| + F_0 |\tau - \tau_0|) \exp (F_1 (t - \varphi)) \\ &\quad \leqq (\tfrac{3}{2} G_0 \tau_0 |\mu| + |h| + F_0 |\tau - \tau_0|) \exp (\tfrac{3}{2} F_1 \tau_0). \end{aligned} \tag{3.20}$$

Applying the last inequality in (3.19), we get

$$|\delta f'_x(t)| \leqq F_2 (\tfrac{3}{2} G_0 \tau_0 |\mu| + |h| + F_0 |\tau - \tau_0|) \exp (\tfrac{3}{2} F_1 \tau_0), \qquad t \in [\varphi, \varphi + \tau]. \tag{3.21}$$

Making use of (3.18), (3.21), (3.7) and (3.9), we get from (3.17) the following inequality:

$$\begin{aligned} |\delta x'_{h_i}(t)| \leqq{} & F_1 \frac{K}{n} |\tau - \tau_0| \\ & + \int_\varphi^t \Big((F_1 + |\mu| G_1) |\delta x'_{h_i}(u)| \\ & \qquad + (F_2 (\tfrac{3}{2} G_0 \tau_0 |\mu| + |h| + F_0 |\tau - \tau_0|) \exp (\tfrac{3}{2} F_1 \tau_0) + |\mu| G_1) \frac{K}{n} \Big)\, du, \\ & \qquad\qquad t \in [\varphi, \varphi + \tau]. \end{aligned}$$

Applying Gronwall's lemma, we get

$$\begin{aligned} &|\delta x'_{h_i}(t)| \\ &\quad \leqq \Big((F_2 (\tfrac{3}{2} G_0 \tau_0 |\mu| + |h| + F_0 |\tau - \tau_0|) \exp (\tfrac{3}{2} F_1 \tau_0) + G_1 |\mu|) \frac{K}{n} (t - \varphi) \\ &\qquad + F_1 \frac{K}{n} |\tau - \tau_0| \Big) \exp [(F_1 + |\mu| G_1 (t - \varphi)]. \end{aligned}$$

Substituting $t = \varphi + \tau$ and taking (3.10) into account, we conclude (3.15) readily.

LEMMA 3.3. *If conditions* (3.10) *and* (3.11) *are satisfied, then*

$$
\begin{aligned}
&|\dot{x}(\varphi+\tau;\varphi,p^0+h,\mu,\tau)-\dot{x}(\tau_0;0,p^0,0,\tau_0)|\\
&\quad=|\dot{x}(\varphi+\tau;\varphi,p^0+h,\mu,\tau)-\dot{p}(\tau_0)|\\
&\quad\leqq F_1(\tfrac{3}{2}G_0\tau_0|\mu|+|h|+F_0|\tau-\tau_0|)\exp(\tfrac{3}{2}F_1\tau_0)+G_0|\mu|.
\end{aligned}
\tag{3.22}
$$

Proof. Substituting $t=\varphi+\tau$ into (3.12), $t=\varphi+\tau_0$ into (3.13) and subtracting the two identities, we obtain

$$
\begin{aligned}
&|\dot{x}(\varphi+\tau;\varphi,p^0+h,\mu,\tau)-\dot{p}(\tau_0)|\\
&=\left|f(x(\varphi+\tau;\varphi,p^0+h,\mu,\tau))-f(p(\tau_0))+\mu g\left(\frac{\varphi+\tau}{\tau},x(\varphi+\tau;\varphi,p^0+h,\mu,\tau)\right)\right|\\
&\leqq F_1|x(\varphi+\tau;\varphi,p^0+h,\mu,\tau)-p(\tau_0)|+|\mu|G_0.
\end{aligned}
$$

Taking into account (3.20) at $t=\varphi+\tau$, we can conclude (3.22).

LEMMA 3.4. *If conditions* (3.10) *and* (3.11) *are satisfied, then*

$$
|x'_\tau(\varphi+\tau;\varphi,p^0+h,\mu,\tau)|\leqq|\mu|\frac{2\varphi+3\tau_0}{\tau_0}G_s\exp[\tfrac{3}{2}\tau_0(F_1+G_1|\mu|)].
\tag{3.23}
$$

Proof. Differentiating (3.12) with respect to τ, we get

$$
\begin{aligned}
\dot{x}'_\tau(t;\varphi,p^0+h,\mu,\tau)\equiv\Big(&f'_x(x(t;\varphi,p^0+h,\mu,\tau))\\
&+\mu g'_x\left(\frac{t}{\tau},x(t;\varphi,p^0+h,\mu,\tau)\right)\Big)x'_\tau(t;\varphi,p^0+h,\mu,\tau)\\
&-\mu\frac{t}{\tau^2}g'_s\left(\frac{t}{\tau},x(t;\varphi,p^0+h,\mu,\tau)\right).
\end{aligned}
$$

Integrating from φ to $t>\varphi$, we obtain

$$
\begin{aligned}
&x'_\tau(t;\varphi,p^0+h,\mu,\tau)-x'_\tau(\varphi;\varphi,p^0+h,\mu,\tau)\\
&\equiv\int_\varphi^t\Big(\Big(f'_x(x(u;\varphi,p^0+h,\mu,\tau))+\mu g'_x\left(\frac{u}{\tau},x(u;\varphi,p^0+h,\mu,\tau)\right)\Big)x'_\tau(u;\varphi,p^0+h,\mu,\tau)\\
&\qquad-\mu\frac{u}{\tau^2}g'_s\left(\frac{u}{\tau},x(u;\varphi,p^0+h,\mu,\tau)\right)\Big)\,du.
\end{aligned}
$$

The second term on the left hand side is zero since $x(\varphi;\varphi,p^0+h,\mu,\tau)\equiv p^0+h$ does not depend on τ. Taking into account (3.7) and (3.10) for $t\in[\varphi,\varphi+\tau]$, we have the inequality

$$
|x'_\tau(t;\varphi,p^0+h,\mu,\tau)|\leqq\int_\varphi^t\left((F_1+|\mu|G_1)|x'_\tau(u;\varphi,p^0+h,\mu,\tau)|+|\mu|\frac{\varphi+\tau}{\tau^2}G_s\right)du.
$$

Applying Gronwall's lemma, we have

$$
|x'_\tau(u;\varphi,p^0+h,\mu,\tau)|\leqq|\mu|\frac{\varphi+\tau}{\tau^2}G_s\tau\exp[(F_1+|\mu|G_1)\tau].
$$

$$t\in[\varphi,\varphi+\tau].$$

With (3.10) taken into account again and with the substitution $t=\varphi+\tau$ into the last inequality, (3.23) follows.

The following two lemmata are needed for giving estimates for Δ of (2.3) in our case.

LEMMA 3.5. *If conditions* (3.10) *and* (3.11) *are satisfied, then*

$$\begin{aligned}|x'_\mu(\varphi+\tau;\varphi,p^0+h,\mu,\tau)-x'_\mu(\tau_0;0,p^0,0,\tau_0)|\\ \leqq (\tfrac{3}{2}G_s(|\tau-\tau_0|+|\varphi|)\\ +\tfrac{3}{2}(G_1\tau_0+F_2KK_{-1}G_0\tau_0^2)(\tfrac{3}{2}G_0\tau_0|\mu|+|h|+F_0|\tau-\tau_0|)\exp(\tfrac{3}{2}F_1\tau_0)\\ +\tfrac{3}{2}G_1KK_{-1}G_0\tau_0^2|\mu|+KK_{-1}G_0|\tau-\tau_0|)\exp[\tfrac{3}{2}\tau_0(F_1+G_1|\mu|)].\end{aligned} \tag{3.24}$$

Proof. The proof is similar to the proof of Lemma 3.2 and will not be given in detail. However, attention is drawn to the fact that $x'_\mu(t;0,p^0,0,\tau_0)$ satisfies the inhomogeneous linear system

$$\dot y=f'_x(p(t))y+g\left(\frac{t}{\tau_0},p(t)\right)$$

and $x'_\mu(0;0,p^0,0,\tau_0)=0$, thus

$$x'_\mu(t;0,p^0,0,\tau_0)=Y(t)\int_0^t Y^{-1}(u)g\left(\frac{u}{\tau_0},p(u)\right)du.$$

LEMMA 3.6. *If conditions* (3.10) *and* (3.11) *are satisfied, then*

$$\begin{aligned}|x'_\varphi(\varphi+\tau;\varphi,p^0+h,\mu,\tau)-x'_\varphi(\tau_0;0,p^0,0,\tau_0)|\\ \leqq (\tfrac{3}{2}(\tfrac{3}{2}G_0\tau_0|\mu|+|h|+F_0|\tau-\tau_0|)F_2\tau_0\exp(\tfrac{3}{2}F_1\tau_0)\\ +\tfrac{3}{2}G_1\tau_0|\mu|+F_1|\tau-\tau_0|)P\exp[\tfrac{3}{2}\tau_0(F_1+G_1|\mu|)].\end{aligned} \tag{3.25}$$

Proof. The proof is similar to that of Lemma 3.2 and will not be given in detail. It is to be noted that $x(t;\varphi,p^0,0,\tau_0)\equiv p(t-\varphi)$ and hence $x'_\varphi(t;\varphi,p^0,0,\tau_0)\equiv -p(t-\varphi)$. Thus $|x'_\varphi(t;\varphi,p^0,0,\tau_0)|\leqq P$ (see (3.9)).

4. Main results. The estimates given in the previous lemmata make the application of Ostrowski's results possible to the problem outlined at the beginning of § 3.

Lewis' implicit function theorem (Theorem 2.A) is to be applied to the periodicity condition (3.5) from which $(v_1,\cdots,v_n)=(\tau,h_2,\cdots,h_n)$ are to be expressed as functions of $(u_1,u_2)=(\mu,\varphi)$. The Jacobi matrix $J(\mu,\varphi,\tau,h)$ of z with respect to the variables τ, h at $\mu=0$, $\varphi=0$, $\tau=\tau_0$, $h=0$ is, by (3.6), nonzero. As it is easily seen from the remark preceding (3.9),

$$J(0,0,\tau_0,0)=\begin{bmatrix}\dot p_1^0 & 0 & 0 & \cdots & 0\\ 0 & 0 & 0 & \cdots & 0\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ 0 & 0 & 0 & \cdots & 0\end{bmatrix}+Y(\tau_0)-I$$

where the first term on the right hand side is an n by n matrix whose elements are all zero except the one standing in the first row and first column. This matrix is considered to be known along with its inverse, which will be denoted by

$$J^{-1}=J^{-1}(0,0,\tau_0,0)$$

for short.

Let $a_1, a_2, a_3, \beta_1, \beta_2, \beta_3$ be arbitrary positive numbers satisfying the following inequalities, respectively:

$$\tfrac{3}{2}G_0\tau_0 a_1+\beta_1<\sigma \exp(-\tfrac{3}{2}F_1\tau_0), \tag{4.1}$$

$$\begin{aligned} &a_2(\tfrac{3}{2}F_1G_0\tau_0 \exp(\tfrac{3}{2}F_1\tau_0)+G_0+4G_s \exp[\tfrac{3}{2}\tau_0(F_1+G_1a_2)]) \\ &\qquad +\beta_2(1+F_0)F_1 \exp(\tfrac{3}{2}F_1\tau_0)<\frac{1}{n|J^{-1}|}, \end{aligned} \tag{4.2}$$

$$\begin{aligned} &(\tfrac{3}{2}a_3(\tfrac{3}{2}F_2G_0\tau_0^2 \exp(\tfrac{3}{2}F_1\tau_0)+G_1\tau_0) \\ &\qquad +\beta_3(\tfrac{3}{2}(1+F_0)F_2\tau_0 \exp(\tfrac{3}{2}F_1\tau_0)+F_1)) \exp(\tfrac{3}{2}\tau_0G_1a_3) \\ &\qquad <\frac{1}{|J^{-1}|K} \exp(-\tfrac{3}{2}F_1\tau_0). \end{aligned} \tag{4.3}$$

Further, let us define

$$a=\min\left(a_1, a_2, a_3, \frac{\tau_0}{2}\right),$$

$$\beta=\min\left(\beta_1, \beta_2, \beta_3, \frac{\tau_0}{2}\right),$$

$$\begin{aligned} H_1=n(&a(\tfrac{3}{2}F_1G_0\tau_0 \exp(\tfrac{3}{2}F_1\tau_0)+G_0+4G_s \exp[\tfrac{3}{2}\tau_0(F_1+G_1a)]) \\ &+\beta(1+F_0)F_1 \exp(\tfrac{3}{2}F_1\tau_0)), \end{aligned}$$

$$\begin{aligned} H_2=(&\tfrac{3}{2}a(\tfrac{3}{2}F_2G_0\tau_0^2 \exp(\tfrac{3}{2}F_1\tau_0)+G_1\tau_0) \\ &+\beta(\tfrac{3}{2}(1+F_0)F_2\tau_0 \exp(\tfrac{3}{2}F_1\tau_0)+F_1))K \exp[\tfrac{3}{2}\tau_0(F_1+G_1a)], \end{aligned}$$

$$H=\max(H_1, H_2),$$

$$\begin{aligned} \Delta_1=|J^{-1}|KK_{-1}G_0\tau_0+\frac{|J^{-1}|}{1-|J^{-1}|H}(&(\tfrac{3}{2}G_s(a+\beta) \\ &+\tfrac{3}{2}(G_1\tau_0+F_2KK_{-1}G_0\tau_0^2)(\tfrac{3}{2}G_0\tau_0a+\beta+F_0\beta) \exp(\tfrac{3}{2}F_1\tau_0) \\ &+\tfrac{3}{2}G_1KK_{-1}G_0\tau_0^2a+KK_{-1}G_0\beta) \exp[\tfrac{3}{2}\tau_0(F_1+G_1a)] \\ &+|J^{-1}|HKK_{-1}G_0\tau_0), \end{aligned}$$

$$\begin{aligned} \Delta_2=\frac{|J^{-1}|}{1-|J^{-1}|H}(&F_1(\tfrac{3}{2}G_0\tau_0a+\beta+F_0\beta) \exp(\tfrac{3}{2}F_1\tau_0) \\ &+G_0a+(\tfrac{3}{2}(\tfrac{3}{2}G_0\tau_0a+\beta+F_0\beta)F_2\tau_0 \exp(\tfrac{3}{2}F_1\tau_0) \\ &+\tfrac{3}{2}G_1\tau_0a+F_1\beta)P \exp[\tfrac{3}{2}\tau_0(F_1+G_1a)]), \end{aligned}$$

$$\Delta=2\max(\Delta_1, \Delta_2). \tag{4.4}$$

Finally we introduce the notation

$$U=\left\{(\mu, \varphi)\in R^2: |\mu|<\min\left(a, \frac{\beta}{\Delta}\right), |\varphi|<\min\left(a, \frac{\beta}{\Delta}\right)\right\}.$$

THEOREM 4.1. *To all* $(\mu, \varphi) \in U$ *there belongs one and only one* $\tau(\mu, \varphi) \in R$ *and one and only one* $h(\mu, \varphi) = (0, h_2(\mu, \varphi), \cdots, h_n(\mu, \varphi))$ *such that the solution of system*

$$\dot{x} = f(x) + g\left(\frac{t}{\tau(\mu, \varphi)}, x\right)$$

which at $t = \varphi$ *assumes the value* $x = p^0 + h(\mu, \varphi)$ *is periodic with period* $\tau(\mu, \varphi)$, *the functions* $\tau: U \to R$, $h: U \to R^n$ *having the properties* $\tau \in C_U^1$, $h \in C_U^1$, $|\tau(\mu, \varphi) - \tau_0| < \beta$, $|h(\mu, \varphi)| < \beta$ *in* U, $\tau(0, 0) = \tau_0$, $h(0, 0) = 0$.

Proof. The proof consists of checking the validity of Theorem 2.A for the system of equations (3.5).

We have for the variation of the Jacobi matrix of the system

$$\begin{aligned} &|J(\mu, \varphi, \tau, h) - J(0, 0, \tau_0, 0)| \\ &\quad = n \max \{|\dot{x}(\varphi + \tau; \varphi, p^0 + h, \mu, \tau) - \dot{p}(\tau_0)| + |x'_\tau(\varphi + \tau; \varphi, p^0 + h, \mu, \tau)|, \\ &\qquad |x'_{h_i}(\varphi + \tau; \varphi, p^0 + h, \mu, \tau) - x'_{h_i}(\tau_0; 0, p^0, 0, \tau_0)|, (i = 2, 3, \cdots, n)\}. \end{aligned}$$

Applying Lemmata 3.3, 3.4 and 3.2 and taking into consideration (4.2) and (4.3), we see that for $|\mu| \leqq a$, $|\varphi| \leqq a$, $|\tau - \tau_0| \leqq \beta$ and $|h| \leqq \beta$,

$$|J(\mu, \varphi, \tau, h) - J(0, 0, \tau_0, 0)| \leqq H < \frac{1}{|J^{-1}|}.$$

Thus, by (3.6) and Theorem 2.B, the matrix $J(\mu, \varphi, \tau, h)$ is regular.

We have two systems now for (2.2), namely,

$$J(\mu, \varphi, \tau, h) d^1 = -x'_\mu(\varphi + \tau; \varphi, p^0 + h, \mu, \tau) \tag{4.5}$$

and

$$J(\mu, \varphi, \tau, h) d^2 = -(\dot{x}(\varphi + \tau; \varphi, p^0 + h, \mu, \tau) + x'_\varphi(\varphi + \tau; \varphi, p^0 + h, \mu, \tau)). \tag{4.6}$$

As is seen from the proof of Lemma 3.5,

$$x'_\mu(\tau_0; 0, p^0, 0, \tau_0) = Y(\tau_0) \int_0^{\tau_0} Y^{-1}(t) g\left(\frac{t}{\tau_0}, p(t)\right) dt.$$

Clearly, $\dot{x}(\tau_0; 0, p^0, 0, \tau_0) = \dot{p}(\tau_0) = \dot{p}(0) = \dot{p}^0$, and since $x(t; \varphi, p^0, 0, \tau_0) \equiv p(t - \varphi)$, we have $x'_\varphi(\tau_0; 0, p^0, 0, \tau_0) = -\dot{p}^0$.

Therefore at $\mu = \varphi = 0$, $\tau = \tau_0$ and $h = 0$, systems (4.5) and (4.6) assume the form

$$J(0, 0, \tau_0, 0) d^1 = -Y(\tau_0) \int_0^{\tau_0} Y^{-1}(t) g\left(\frac{t}{\tau_0}, p(t)\right) dt,$$

$$J(0, 0, \tau_0, 0) d^2 = 0,$$

respectively. Hence

$$d^1 = -J^{-1} Y(\tau_0) \int_0^{\tau_0} Y^{-1}(t) g\left(\frac{t}{\tau_0}, p(t)\right) dt, \qquad d^2 = 0.$$

Estimates can be given for $d^1(\mu, \varphi, \tau, h)$ and $d^2(\mu, \varphi, \tau, h)$ and thus for the n by 2 matrix $D(\mu, \varphi, \tau, h) = [d^1, d^2]$ by applying Lemmata 3.5, 3.3, 3.6 and Theorem 2.B. It turns out that for $|\mu| \leqq a$, $|\varphi| \leqq a$, $|\tau - \tau_0| \leqq \beta$, $|h| \leqq \beta$ we have $|d^1(\mu, \varphi, \tau, h)| \leqq \Delta_1$, $|d^2(\mu, \varphi, \tau, h)| \leqq \Delta_2$ and, as a consequence,

$$|D(\mu, \varphi, \tau, h)| \leqq \Delta$$

where Δ has been defined by (4.4). Thus condition (2.3) is fulfilled with our Δ, and this completes the proof.

Attention is drawn to the fact that this result contains two important special cases. The first case occurs when the perturbation does not depend on the state of the system, i.e. in system (3.1) the function g does not depend on x:

$$\dot{x} = f(x) + \mu g\left(\frac{t}{\tau}\right). \tag{4.7}$$

The rest of the assumptions made for system (3.1) and the unperturbed system (3.2) are considered to remain valid. In this case $G_1 = 0$ and, as a consequence, inequalities (4.2) and (4.3) for the determination of $a_2, a_3, \beta_2, \beta_3$ become linear. The determination of a, β and also of Δ becomes considerably simpler.

The second case occurs when the perturbation is autonomous, i.e. the function g does not depend on t:

$$\dot{x} = f(x) + \mu g(x), \tag{4.8}$$

while the rest of the assumptions remain valid. In this case $G_s = 0$ and, as a consequence, inequality (4.2) is linear and the expressions for H_1 and Δ_1 are simpler. Since the perturbed system (4.8) is autonomous the functions $\tau: U \to R$ and $h: U \to R^n$ are constant in the variable φ, i.e. the period and the initial value of the perturbed periodic solution depend solely on $\mu \in (-a, a) \cap (-\beta/\Delta, \beta/\Delta)$ and not on the "initial phase".

REFERENCES

[1] M. FARKAS, *Controllably periodic perturbations of autonomous systems*, Acta Math. Acad. Sci. Hungar., 22 (1971), pp. 337–348.

[2] ———, *Determination of controllably periodic perturbed solutions by Poincaré's method*, Studia Sci. Math. Hungar., 7 (1972), pp. 257–266.

[3] ———, *On isolated periodic solutions of differential systems*, Ann. Mat. Pura Appl., 106 (1975), pp. 233–243.

[4] H. I. FREEDMAN, *Estimates on the existence region for periodic solutions of equations involving a small parameter. I: The noncritical case*, SIAM J. Appl. Math., 16 (1968), pp. 1341–1349.

[5] ———, *Estimates on the existence region for periodic solutions of equations involving a small parameter. II: Critical cases*, Ann. Mat. Pura Appl., 90 (1971), pp. 259–279.

[6] D. C. LEWIS, *Invariant manifolds near an invariant point of unstable type*, Amer. J. Math., 60 (1938), pp. 577–587.

[7] ———, *Periodic solutions of differential equations containing a parameter*, Duke Math. J., 22 (1955), pp. 39–56.

[8] A. OSTROWSKI, *Sur la détermination des bornes inférieures pour une classe des déterminants*, Bull. Sci. Math., 61 (1937), No. I, pp. 19–32.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0013 $01.00/0

A POSITIVE KERNEL FOR HAHN-EBERLEIN POLYNOMIALS*

MIZAN RAHMAN†

Abstract. Explicit forms of the coefficients $E(x, y, z)$ in the expansion $Q_n(x)Q_n(y) = \sum_{z=0}^{N} E(x, y, z)Q_n(z)$, where $Q_n(x) = Q_n(x; \alpha, \beta, N)$ is the Hahn polynomial in the integer-valued variable x, $0 \leqq x \leqq N$, are given. It is shown that if $\alpha \leqq \beta < -N$ and $\beta - \alpha$ is a nonnegative integer, then E is nonnegative for all x, y, z and N. However, if $\alpha \leqq \beta < -N$ but $\beta - \alpha$ is nonintegral, then E is nonnegative only under the stronger restriction $\beta - \alpha > N - 1$.

1. Introduction. Positivity of integrals and sums of special functions has been a theme of many recent papers (see, for example, [1]–[9], [14]–[20], [22]–[25]; for an excellent review article on this topic and for further references see Gasper [20]). Apart from their well-known applications in harmonic analysis, certain positive sums and integrals have found applications in coding theory [27] and probability theory [11], [21]. The positive sum that we are going to present here may indeed be viewed as an extension of certain results known in coding theory mainly through the works of Delsarte [12].

In [15] Gasper studied a convolution structure for the Jacobi series by computing the kernel $K(x, y, z; \alpha, \beta)$ where

$$R_n^{(\alpha,\beta)}(x)R_n^{(\alpha,\beta)}(y) = \int_{-1}^{1} K(x, y, z; \alpha, \beta)R_n^{(\alpha,\beta)}(z)(1-z)^\alpha(1+z)^\beta\, dz, \tag{1.1}$$

with $R_n^{(\alpha,\beta)}(x) = P_n^{(\alpha,\beta)}(x)/P_n^{(\alpha,\beta)}(1)$, and then establishing the nonnegativity of the kernel for $\alpha \geqq \beta \geqq -\frac{1}{2}$. Since the Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ are known to be the continuous limit of the Hahn polynomials

$$Q_n(x) = Q_n(x; \alpha, \beta, N) = {}_3F_2\left[\begin{matrix} -n, & n+\alpha+\beta+1, & -x \\ & \alpha+1, & -N \end{matrix}\right], \tag{1.2}$$

$x = 0, 1, \cdots, N$; $n = 0, 1, \cdots, N$, a similar convolution structure can be expected for these polynomials as well. However, as Gasper observed in [20, p. 393] the coefficients in

$$Q_n(x; \alpha, \beta, N)Q_n(y; \alpha, \beta, N) = \sum_{z=0}^{N} E(x, y, z; \alpha, \beta, N)Q_n(z; \alpha, \beta, N) \tag{1.3}$$

are not all nonnegative for $\alpha, \beta > -1$. So we need to consider the case when either α or β or both are less than or equal to -1. In particular, if $\alpha, \beta < -N$ the ${}_3F_2$ functions on the right of (1.2) remain well-defined, and are orthogonal with a positive weight function. In this region there appears to be no continuous analogue of the Hahn polynomials and, in fact, they may be called Hahn–Eberlein polynomials. The Eberlein polynomials [13] used by Delsarte [12] and Sloane [27] correspond to the case

$$\beta = -N-1, \qquad \alpha = -V+N-1 \quad \text{with } V \geqq 2N. \tag{1.4}$$

When α and β are restricted by (1.4), a nonnegative coefficient in the expansion (1.3) has been found by algebraic methods in coding theory (see, for example, Sloane [27, p. 244]). The purpose of this present paper is to approach the problem from hypergeometric series point of view and, of course, to establish some general criteria for the nonnegativity of the coefficients $E(x, y, z; \alpha, \beta, N)$ in (1.3).

* Received by the editors May, 1976, and in final revised form February 23, 1977.

† Department of Mathematics, Carleton University, Ottawa, Ontario, Canada K1S 5B6. This work was supported by the National Research Council of Canada under Grant A 6197.

Apart from some well-known formulas of hypergeometric series [10], [26] two important tools in our work are the following formulas:

$$Q_n(x;\alpha,\beta,N)Q_n(y;\alpha,\beta,N)=\frac{(-1)^n(\beta+1)_n}{(\alpha+1)_n}\sum_{r=0}^{n}\sum_{s=0}^{n-r}\frac{(-n)_{r+s}(n+\alpha+\beta+1)_{r+s}}{(-N)_{r+s}(-N)_{r+s}}\cdot\frac{(-x)_r(-y)_r(x-N)_s(y-N)_s}{(\alpha+1)_r(\beta+1)_s r!s!} \tag{1.5}$$

[17]

$$(1.6)\quad {}_{p+1}F_{q+1}\begin{bmatrix}-x & a_1,\cdots,a_p\\ c+\mu & b_1,\cdots,b_q\end{bmatrix}=\sum_{y=0}^{x}\binom{x}{y}\frac{(c)_y(\mu)_{x-y}}{(c+\mu)_x}\,{}_{p+1}F_{q+1}\begin{bmatrix}-y, & a_1,\cdots,a_p\\ c, & b_1,\cdots,b_q\end{bmatrix},$$

[18].

We have also made extensive use of the following identities in Pochhammer products [26, p. 239]:

$$(a)_{m+n}=(a)_m(a+m)_n,\quad (a)_{m-n}=(-1)^n(a)_m/(1-a-m)_n,$$
$$(a)_m=(-1)^m(1-a-m)_m.$$

Using orthogonality of the Hahn polynomials with respect to the weight function

$$\rho(x;\alpha,\beta,N)=\binom{x+\alpha}{x}\binom{N-x+\beta}{N-x}\Big/\binom{N+\alpha+\beta+1}{N}, \tag{1.7}$$

we can invert (1.3) to obtain

$$K(x,y,z;\alpha,\beta,N)=\sum_{n=0}^{N}\pi_nQ_n(x;\alpha,\beta,N)Q_n(y;\alpha,\beta,N)Q_n(z;\alpha,\beta,N), \tag{1.8}$$

where

$$K(x,y,z;\alpha,\beta,N)=E(x,y,z;\alpha,\beta,N)/\rho(z;\alpha,\beta,N)$$

and

$$\pi_n=\frac{(-1)^n(-N)_n(\alpha+1)_n(\alpha+\beta+1)_n}{n!(\beta+1)_n(N+\alpha+\beta+2)_n}\cdot\frac{2n+\alpha+\beta+1}{\alpha+\beta+1}. \tag{1.9}$$

We may now discuss the case in which either α or β or both are between $-N$ and -1.

Obviously neither α nor β can be an integer in that range since $\rho(x)$, $Q_n(x)$ and π_n could become infinite. Even for nonintegral α,β we can show that not all E can be positive. This can be settled here if we quote (2.13) and (2.14), which we derive in the next section, for the special case $x+z=y$. A little simplification yields

$$E(x,y,y-x;\alpha,\beta,N)=K(x,y,y-x;\alpha,\beta,N)\rho(y-x;\alpha,\beta,N)=\frac{\binom{y}{x}}{\binom{N}{x}}\frac{(\beta+1+N-y)_x}{(\alpha+1)_x}. \tag{1.10}$$

In particular, if we set $x=1$, $\alpha=-\frac{5}{2}$, $\beta=-\frac{3}{2}$ we get

$$E(1,y,y-1;-\tfrac{5}{2},-\tfrac{3}{2},N)=-\frac{2y}{3N}(N-y-\tfrac{1}{2}).$$

This is negative if $N>y$ and positive if $N=y$.

Also setting $\alpha=-N+\frac{3}{2}$, $\beta=-N+\frac{1}{2}$ we obtain

$$E(x, y, y-x; -N+\tfrac{3}{2}, -N+\tfrac{1}{2}, N)=\frac{\binom{y}{x}}{\binom{N}{x}}\cdot\frac{(\frac{3}{2}-y)_x}{(-N+\frac{5}{2})_x}.$$

Again the sign of E depends on N, x and y. For instance if we take $x=1$, $N=y=2$ then $E=-1$, while for $x=1$, $y=2$, $N=3$ we get $E=\frac{2}{3}$. The situation remains essentially the same when only one of the parameters α, β lies between $-N$ and -1. Suppose $-N\leqq\beta\leqq-1$ and α is arbitrary. Writing $t=\beta+N$, $0\leqq t\leqq N-1$, the factor that determines the sign of E in (1.10) is $(t+1-y)_x/(\alpha+1)_x$. For $x\geqq 1$, this blows up at $\alpha=-1$ if $y\neq t+1$, and is nonnegative if $\alpha>-1$ and $y\leqq t+1$ but negative if x is odd and $y>t+x$. On the other hand, if $\alpha<-N$ then $(\alpha+1)_x$ has the sign $(-1)^x$ but the numerator $(t+1-y)_x$ can be of any sign depending on the magnitude of t and y.

The search for a nonnegative representation of E can, therefore, be concentrated on the region $\alpha, \beta<-N$. In § 2 we shall obtain a double series representation of $K(x, y, z)\equiv K(x, y, z; \alpha, \beta, N)$ and in § 3 and § 4 we shall obtain conditions under which this is nonnegative.

2. A double series representation of $K(x, y, z)$. Our first step is to observe that the double sum in (1.5) can be replaced by

$$Q_n(x;\alpha,\beta,N)Q_n(y;\alpha,\beta,N)$$
$$=\frac{(-1)^n(\beta+1)_n}{(\alpha+1)_n}\sum_{r=0}^{N}\sum_{s=0}^{N-r}\frac{(-n)_{r+s}(n+\alpha+\beta+1)_{r+s}(-x)_r(-y)_r(x-N)_s(y-N)_s}{(-N)_{r+s}(-N)_{r+s}(\alpha+1)_r(\beta+1)_s r!s!}, \tag{2.1}$$

$x, y, n=0, 1, \cdots, N$; likewise we shall write

$$Q_n(z;\alpha,\beta,N)=\sum_{k=0}^{N}\frac{(-n)_k(-z)_k(n+\alpha+\beta+1)_k}{(\alpha+1)_k(-N)_k k!}, \tag{2.2}$$

$z=0, 1, \cdots, N$.

Substituting (2.1) and (2.2) in (1.8), we obtain

$$K(x, y, z)=\sum_{k=0}^{N}\frac{(-z)_k}{(\alpha+1)_k(-N)_k k!}\sum_{r=0}^{N}\sum_{s=0}^{N-r}\cdot\frac{(-x)_r(-y)_r(x-N)_s(y-N)_s}{(-N)_{r+s}(-N)_{r+s}(\alpha+1)_r(\beta+1)_s r!s!}T_{r+s,k} \tag{2.3}$$

where

$$T_{p,k}=\sum_{n=0}^{N}\frac{(-N)_n(-n)_k(-n)_p(\alpha+\beta+1)_n(\alpha+\beta+1+n)_k(\alpha+\beta+1+n)_p}{n!(N+\alpha+\beta+2)_n}\cdot\frac{2n+\alpha+\beta+1}{\alpha+\beta+1}$$
$$=\sum_{n=\max(p,k)}^{N}\frac{(-N)_n(-n)_k(-n)_p(\alpha+\beta+1)_n(\alpha+\beta+1+n)_k(\alpha+\beta+1+n)_p}{n!(N+\alpha+\beta+2)_n}\cdot\frac{2n+\alpha+\beta+1}{\alpha+\beta+1}. \tag{2.4}$$

Evidently the sum on the right is symmetric in p, k so we need to consider only one case, say, $k \geqq p$. Then

$$T_{p,k} = (-1)^{p+k} \sum_{n=k}^{N} \frac{(-N)_n n!(\alpha+\beta+1)_n(\alpha+\beta+1+n)_p(\alpha+\beta+1+n)_k}{(n-p)!(n-k)!(N+\alpha+\beta+2)_n} \cdot \frac{2n+\alpha+\beta+1}{\alpha+\beta+1}, \tag{2.5}$$

which can be shown to reduce to a very well-poised hypergeometric series

$$T_{p,k} = \frac{(-1)^p N!k!(\alpha+\beta+1)_{p+k}(\alpha+\beta+2)_{2k}}{(\alpha+\beta+1)_k(N+\alpha+\beta+2)_k(N-k)!(k-p)!} \cdot {}_5F_4\left[\begin{matrix}\alpha+\beta+1+2k, & \frac{\alpha+\beta+1}{2}+k+1, & \alpha+\beta+1+k+p, & k+1, & -N+k \\ & \frac{\alpha+\beta+1}{2}+k, & \alpha+\beta+1+k, & k-p+1, & N+\alpha+\beta+2+k\end{matrix}\right] \tag{2.6}$$

This series can be summed by using a special case of Dougall's theorem for very well-poised hypergeometric series (see, for example, Slater [26, p. 56]):

$$ {}_5F_4\left[\begin{matrix}a & 1+\frac{1}{2}a, & b, & c, & -m \\ & \frac{1}{2}a, & 1+a-b, & 1+a-c, & 1+a+m\end{matrix}\right] = \frac{(1+a)_m(1+a-b-c)_m}{(1+a-b)_m(1+a-c)_m}, \quad m=0,1,2,\cdots. \tag{2.7}$$

We thus obtain

$$T_{p,k} = \begin{cases} 0 & \text{if } k+p<N, \\ (-1)^{k+p+N} \dfrac{N!(\alpha+\beta+2)_N}{(\alpha+\beta+1)_N} \cdot \dfrac{k!p!(\alpha+\beta+1)_{k+p}}{(N-k)!(N-p)!(k+p-N)!} & \text{if } k+p \geqq N. \end{cases} \tag{2.8}$$

Hence

$$K(x,y,z) = \frac{(-1)^N(\alpha+\beta+2)_N}{N!(\alpha+\beta+1)_N} \sum_{r=0}^{N} \sum_{s=0}^{N-r} \frac{(-x)_r(-y)_r(x-N)_s(y-N)_s(r+s)!}{(-N)_{r+s}(\alpha+1)_r(\beta+1)_s r!s!} \cdot \sum_{k=N-r-s}^{N} \frac{(-z)_k(\alpha+\beta+1)_{k+r+s}}{(\alpha+1)_k(k+r+s-N)!},$$

provided $z \geqq N-r-s \geqq 0$, and 0 otherwise.

Using the Chu–Vandermonde theorem [5]

$$ {}_2F_1\left[\begin{matrix}-n, & a \\ & b\end{matrix}\right] = \frac{(b-a)_n}{(b)_n}, \tag{2.9}$$

we then have

$$K(x,y,z) = \frac{(\alpha+\beta+2)_N}{N!(\alpha+1)_z(\beta+1)_{N-z}} \sum_{\substack{r=0 \\ r+s\geqq N-z}}^{\min(x,y)} \sum_{s=0}^{N-r} (-1)^{z+r+s} \cdot \frac{(-x)_r(-y)_r(x-N)_s(y-N)_s(-z)_{N-r-s}(r+s)!(\beta+1)_{r+s}}{(-N)_{r+s}(\alpha+1)_r(\beta+1)_s r!s!}. \tag{2.10}$$

For the sake of definiteness let us assume

$$z \leqq x \leqq y. \tag{2.11}$$

Symmetry of the kernel will then allow us to extend the results to the cases with x, y, z interchanged.

For further reduction of $K(x, y, z)$ two cases must be distinguished:

$$\text{(i) } x+z \leqq N; \qquad \text{(ii) } x+z>N.$$

In case (i) it can be seen that

$$K(x, y, z)=0 \quad \text{if } x+z<y \tag{2.12}$$

and

$$K(x, y, z)=\left(\frac{z!}{N!}\right)^2 \frac{(N-z)!(\alpha+\beta+2)_N}{(\alpha+1)_z} \sum_{r=y-z}^{x} \frac{(-x)_r(-y)_r(x-N)_{N-z-r}(y-N)_{N-z-r}}{(\alpha+1)_r(\beta+1)_{N-z-r}(N-z-r)!r!}$$
$$\cdot \sum_{s=0}^{z-y+r} \frac{(-z+y-r)_s(-z+x-r)_s(N-z+1)_s(N-z+\beta+1)_s}{(-z)_s(N-z+\beta+1-r)_s(N-z+1-r)_s s!}$$

if $x+z \geqq y$.

Replacing $r-y+z$ by k, we get

$$\begin{aligned} K(x, y, z)=A &\sum_{k=0}^{x+z-y} \frac{(-x-z+y)_k(-N+y)_k(-z)_k(-\beta-N+y)_k}{(1+y-x)_k(1+y-z)_k(\alpha+1+y-z)_k k!} \\ &\cdot {}_4F_3\begin{bmatrix} -k-y+x & N-z+1, & N-z+\beta+1, & -k \\ -z, & N-y+1-k, & N-y+\beta+1-k & \end{bmatrix}, \end{aligned} \tag{2.13}$$

where

$$A=\frac{x!y!z!(N-x)!(N-z)!(\alpha+\beta+2)_N}{(N!)^2(x+z-y)!(y-x)!(y-z)!(\alpha+1)_{y-z}(\beta+1)_{N-y}(\alpha+1)_z}. \tag{2.14}$$

Case (ii) is a little less straightforward than case (i). One must first split the double sum in (2.10) into two parts as:

$$\sum_{\substack{r=0 \\ r+s\geqq N-z}}^{x} \sum_{s=0}^{N-r} = \sum_{r=0}^{N-z} \sum_{s=N-z-r}^{N-r} + \sum_{r=N-z+1}^{x} \sum_{s=0}^{N-r}.$$

For the first series on the right we replace $s-N+z+r$ by s in (2.10) and for the second series we replace $r-N+z-1$ by r. Thus we obtain

$$K(x, y, z)=K_1(x, y, z)+K_2(x, y, z), \tag{2.15}$$

where

$$\begin{aligned} K_1(x, y, z)=B &\sum_{r=0}^{N-y} \frac{(-N+x)_r(-N+y)_r(-N+z)_r(-\alpha-N+z)_r}{(1+x+z-N)_r(1+y+z-N)_r(\beta+1)_r r!} \\ &\cdot {}_4F_3\begin{bmatrix} -N+y+r, & -N+x+r, & N-z+1, & N-z+\beta+1 \\ -z, & r+1, & \beta+1+r & \end{bmatrix}, \end{aligned} \tag{2.16}$$

$$\begin{aligned} K_2(x, y, z)=B &\sum_{r=0}^{x+z-N-1} \frac{(-x-z+N)_{r+1}(-y-z+N)_{r+1}(N-z+\beta+1)_{r+1}}{(-z)_{r+1}(N-z+\alpha+1)_{r+1}(r+1)!} \\ &\cdot {}_4F_3\begin{bmatrix} -N+y, & -N+x, & N-z+r+2, & N-z+r+\beta+2 \\ -z+1+r, & r+2, & \beta+1 & \end{bmatrix}, \end{aligned} \tag{2.17}$$

with

$$B=\left(\frac{z!}{N!}\right)^2\frac{x!y!(\alpha+\beta+2)_N}{(x+z-N)!(y+z-N)!(\alpha+1)_z(\alpha+1)_{N-z}}. \tag{2.18}$$

We shall now combine K_1 and K_2 by using the following reductions:

$$\begin{aligned}
B^{-1}K_1(x,y,z)&=\sum_{r=0}^{N-y}\sum_{k=0}^{N-y-r}\frac{(-N+x)_{r+k}(-N+y)_{k+r}(-N+z)_r(-\alpha-N+z)_r}{(1+x+z-N)_r(1+y+z-N)_r(\beta+1)_{r+k}(r+k)!}\\
&\quad\cdot\frac{(N-z+1)_k(N-z+\beta+1)_k}{(-z)_kk!}\\
&=\sum_{p=0}^{N-y}\frac{(-N+x)_p(-N+y)_p}{(\beta+1)_pp!}\sum_{k=0}^{p}\frac{(-N+z)_{p-k}(-\alpha-N+z)_{p-k}}{(1+x+z-N)_{p-k}(1+y+z-N)_{p-k}}\\
&\quad\cdot\frac{(N-z+1)_k(N-z+\beta+1)_k}{(-z)_kk!}\\
&=\sum_{p=0}^{N-y}\frac{(-N+x)_p(-N+y)_p(-N+z)_p(-\alpha-N+z)_p}{(1+x+z-N)_p(1+y+z-N)_p(\beta+1)_pp!}\\
&\quad\cdot\sum_{k=0}^{p}\frac{(-x-z+N-p)_k(-y-z+N-p)_k(N-z+1)_k(N-z+\beta+1)_k}{(\alpha+1+N-z-p)_k(N-z+1-p)_k(-z)_kk!}.
\end{aligned} \tag{2.19}$$

Also

$$\begin{aligned}
B^{-1}K_2(x,y,z)&=\sum_{p=0}^{N-y}\frac{(-N+x)_p(-N+y)_p}{(\beta+1)_pp!}\sum_{r=0}^{x+z-N-1}\cdot\frac{(-x-z+N)_{r+1}(-y-z+N)_{r+1}}{(-z)_{r+1}(N-z+\alpha+1)_{r+1}}\cdot\\
&\qquad\frac{(N-z+\beta+1)_{r+1}(N-z+r+2)_p(N-z+r+\beta+2)_p}{(-z+1+r)_p(r+p+1)!}.
\end{aligned}$$

But

$$\begin{aligned}
&\frac{(1+x+z-N)_p(1+y+z-N)_p}{(-\alpha-N+z)_p(-N+z)_p}\cdot\frac{(-x-z+N)_{r+1}(-y-z+N)_{r+1}(N-z+r+2)_p}{(N-z+\alpha+1)_{r+1}}\\
&=\frac{(-x-z+N-p)_p(-x-z+N-p+p)_{r+1}(-y-z+N-p)_p(-y-z+N-p+p)_{r+1}(N-z+1)_{p+r+1}}{(\alpha+1+N-z-p)_p(\alpha+1+N-z-p+p)_{r+1}(1+N-z-p)_p(1+N-z-p+p)_{r+1}}\\
&=\frac{(-x-z+N-p)_{p+r+1}(-y-z+N-p)_{p+r+1}(N-z+1)_{p+r+1}}{(\alpha+1+N-z-p)_{p+r+1}(1+N-z-p)_{p+r+1}}
\end{aligned}$$

Hence

$$\begin{aligned}
B^{-1}K_2(x,y,z)&=\sum_{p=0}^{N-y}\frac{(-N+x)_p(-N+y)_p(-N+z)_p(-\alpha-N+z)_p}{(1+x+z-N)_p(1+y+z-N)_p(\beta+1)_pp!}\\
&\quad\cdot\sum_{r=0}^{x+z-N-1}\frac{(-x-z+N-p)_{p+r+1}(-y-z+N-p)_{p+r+1}(N-z+1)_{p+r+1}(N-z+\beta+1)_{p+r+1}}{(-z)_{p+r+1}(N-z+\alpha+1-p)_{p+r+1}(N-z+1-p)_{p+r+1}(p+r+1)!}.
\end{aligned} \tag{2.20}$$

A sum of the two series on the right of (2.19) and (2.20) can now be written down immediately:

$$
\begin{aligned}
K(x,y,z) &= B \sum_{p=0}^{N-y} \frac{(-N+x)_p(-N+y)_p(-N+z)_p(-\alpha-N+z)_p}{(1+x+z-N)_p(1+y+z-N)_p(\beta+1)_p p!} \\
&\quad \cdot \sum_{k=0}^{x+z+p-N} \frac{(-x-z+N-p)_k(-y-z+N-p)_k(N-z+1)_k(N-z+\beta+1)_k}{(N-z+\alpha+1-p)_k(N-z+1-p)_k(-z)_k k!} \\
&= B \sum_{p=0}^{N-y} \frac{(-N+x)_p(-N+y)_p(-N+z)_p(-\alpha-N+z)_p}{(1+x+z-N)_p(1+y+z-N)_p(\beta+1)_p p!} \\
&\quad \cdot {}_4F_3\left[\begin{matrix}-x-z+N-p, & -y-z+N-p, & N-z+1, & N-z+\beta+1 \\ N-z+\alpha+1-p, & N-z+1-p, & -z & \end{matrix}\right].
\end{aligned} \tag{2.21}
$$

3. Nonnegativity of $K(x,y,z)$ when $x+z \leqq N$. Observing that, by reversing the order of summation, we get

$$
\begin{aligned}
&{}_4F_3\left[\begin{matrix}-k-y+x, & N-z+1, & N-z+\beta+1, & -k \\ -z, & N-y+1-k, & N-y+\beta+1-k & \end{matrix}\right] \\
&= \frac{(-k-y+x)_k(N-z+1)_k(N-z+\beta+1)_k}{(-z)_k(N-y+1-k)_k(N-y+\beta+1-k)_k}(-1)^k \\
&\quad \cdot {}_4F_3\left[\begin{matrix} z+1-k, & -\beta-N+y, & -N+y, & -k \\ 1+y-x, & -N+z-k, & -\beta-N+z-k & \end{matrix}\right],
\end{aligned}
$$

we can reduce the kernel in (2.13) to

$$
K(x,y,z) = A \sum_{k=0}^{x+z-y} \frac{(-x-z+y)_k(N-z+\beta+1)_k}{(\alpha+1+y-z)_k k!} a_k, \tag{3.1}
$$

where

$$
a_k = \frac{(N-z+1)_k}{(y-z+1)_k} {}_4F_3\left[\begin{matrix} -k, & -N+y, & -\beta-N+y, & z+1-k \\ -N+z-k, & 1+y-x, & -\beta-N+z-k & \end{matrix}\right]. \tag{3.2}
$$

It is at this point that formula (1.6) seems crucial. Choosing $a_1 = c = -N+y$ and $\mu = -y+z-k$, we obtain

$$
\begin{aligned}
&{}_4F_3\left[\begin{matrix} -k, & -N+y, & -\beta-N+y, & z+1-k \\ -N+z-k, & 1+y-x, & -\beta-N+z-k & \end{matrix}\right] \\
&= \sum_{l=0}^{k} \binom{k}{l} \frac{(-N+y)_l(-y+z-k)_{k-l}}{(-N+z-k)_k} {}_3F_2\left[\begin{matrix} -l, & -\beta-N+y, & z+1-k \\ 1+y-x, & -\beta-N+z-k & \end{matrix}\right] \\
&= \frac{(1+y-z)_k}{(N-z+1)_k} \sum_{l=0}^{k} \frac{(-k)_l(-N+y)_l}{(1+y-z)_l l!} {}_3F_2\left[\begin{matrix} -l, & -\beta-N+y, & z+1-k \\ 1+y-x, & -\beta-N+z-k & \end{matrix}\right].
\end{aligned} \tag{3.3}
$$

Anticipating the simpler results that follow later, let us set

$$
\beta+N+1 = -\varepsilon. \tag{3.4}
$$

Then (3.1) transforms to

$$K(x,y,z)=A\sum_{l=0}^{x+z-y}\sum_{k=0}^{x+z-y-l}\frac{(-x-z+y)_{k+l}(-z-\varepsilon)_{k+l}(-N+y)_l}{(1+y-z)_l(\alpha+1+y-z)_{k+l}k!\,l!}(-1)^l \cdot {}_3F_2\left[\begin{matrix}-l, & 1+y+\varepsilon, & z+1-k-l\\ 1+y-x, & z+1+\varepsilon-k-l\end{matrix}\right]. \tag{3.5}$$

We now make use of yet another formula, due to Thomae [10, pp. 17–19] but given explicitly in Gasper [18]:

$$ {}_3F_2\left[\begin{matrix}-n, & a, & b\\ c, & d\end{matrix}\right]=\frac{(d-b)_n}{(d)_n}\,{}_3F_2\left[\begin{matrix}-n, & c-a, & b\\ c, & 1+b-d-n\end{matrix}\right]. \tag{3.6}$$

Use of this in the ${}_3F_2$ series above gives

$$ {}_3F_2\left[\begin{matrix}-l, & 1+y+\varepsilon, & z+1-k-l\\ 1+y-x, & z+1+\varepsilon-k-l\end{matrix}\right] =\frac{(-x-\varepsilon)_l}{(1+y-x)_l}\sum_{m=0}^{l}\frac{(-l)_m(\varepsilon)_m(1+y+\varepsilon)_m}{(1+x+\varepsilon-l)_m(1+z+\varepsilon-k-l)_m m!}.$$

Simplifying some Pochhammer products, we then obtain

$$K(x,y,z)=A\sum_{l=0}^{x+z-y}\frac{(-x+z+y)_l(-N+y)_l(-x-\varepsilon)_l}{(1+y-z)_l(1+y-x)_l(\alpha+1+y-z)_l l!}(-1)^l \cdot\sum_{m=0}^{l}\frac{(-l)_m(\varepsilon)_m(1+y+\varepsilon)_m(-z-\varepsilon)_{l-m}}{(1+x+\varepsilon-l)_m m!}(-1)^m \cdot\sum_{k=0}^{x+z-y-l}\frac{(-x-z+y+l)_k(-z-\varepsilon+l-m)_k}{(\alpha+1+y-z+l)_k k!}.$$

But the last series over k sums to

$$\frac{(\alpha+1+y+\varepsilon+m)_{x+z-y-l}}{(\alpha+1+y-z+l)_{x+z-y-l}}$$

by the Chu–Vandermonde theorem (2.9). Hence $K(x,y,z)$ becomes

$$K(x,y,z)=A\frac{(\alpha+1+y+\varepsilon)_{x+z-y}}{(\alpha+1+y-z)_{x+z-y}}\sum_{l=0}^{x+z-y}\frac{(-x-z+y)_l(-N+y)_l(-x-\varepsilon)_l(-z-\varepsilon)_l}{(1+y-x)_l(1+y-z)_l(-\alpha-x-z-\varepsilon)_l l!} \cdot {}_4F_3\left[\begin{matrix}-l, & \varepsilon, & 1+y+\varepsilon, & \alpha+1+x+z+\varepsilon-l\\ 1+x+\varepsilon-l, & 1+z+\varepsilon-l, & \alpha+1+y+\varepsilon\end{matrix}\right]. \tag{3.7}$$

In the special case $\varepsilon=0$ we obtain

$$\begin{aligned}K(x,y,z)&=K(x,y,z;\alpha,-N-1,N)\\ &=A\frac{(\alpha+1+y)_{x+z-y}}{(\alpha+1+y-z)_{x+z-y}}\,{}_4F_3\left[\begin{matrix}-x, & -z, & -N+y, & -x-z+y\\ 1+y-x, & 1+y-z, & -\alpha-x-z\end{matrix}\right].\end{aligned} \tag{3.8}$$

Since we are considering the case $y\leqq x+z\leqq N$, the series on the right of (3.8) is obviously positive if

$$\alpha\leqq -N-1. \tag{3.9}$$

It can be seen easily from (2.13) that the coefficient A is also positive for these values of α and β.

For $\varepsilon \neq 0$, the ${}_4F_3$ series in (3.7) does not have any obvious sign pattern in the case (3.9). Further transformations are therefore called for. Fortunately the sum of the denominator parameters of this ${}_4F_3$ exceeds the sum of the numerator parameters by 1 and hence it is balanced. Using the well-known transformation formula [10; p.56] for a balanced ${}_4F_3$:

(3.10)
$$ {}_4F_3\left[\begin{matrix}-m, & \xi, & \eta, & \zeta \\ u, & v, & w & \end{matrix}\right] = \frac{(v-\zeta)_m(w-\zeta)_m}{(v)_m(w)_m}\,{}_4F_3\left[\begin{matrix} & -m, & u-\xi, & u-\eta, & \zeta \\ u, & 1-v+\zeta-m, & 1-w+\zeta-m & \end{matrix}\right], $$

$m=0,1,2,\cdots;\ u+v+w=\xi+\eta+\zeta-m+1$, we get

$$ {}_4F_3\left[\begin{matrix}-l, & \varepsilon, & 1+y+\varepsilon, & \alpha+1+x+z+\varepsilon-l \\ 1+x+\varepsilon-l, & 1+z+\varepsilon-l, & \alpha+1+y+\varepsilon & \end{matrix}\right] $$
$$ = \frac{(1+x-l)_l(1+z-l)_l}{(1+x+\varepsilon-l)_l(1+z+\varepsilon-l)_l}\,{}_4F_3\left[\begin{matrix}-l, & \alpha, & \varepsilon, & l+y-x-z \\ \alpha+1+y+\varepsilon, & -x, & -z & \end{matrix}\right] $$
$$ = \frac{(1+x-l)_l(1+z-l)_l}{(1+x+\varepsilon-l)_l(1+z+\varepsilon-l)_l} $$
$$ \cdot\frac{(z-y-l)_l(x-y-l)_l}{(-x)_l(-z)_l}\,{}_4F_3\left[\begin{matrix}-l, & y+1+\varepsilon, & \alpha+1+y, & l+y-x-z \\ \alpha+1+y+\varepsilon, & 1+y-z, & 1+y-x & \end{matrix}\right] $$
$$ = \frac{(1+y-x)_l(1+y-z)_l}{(-x-\varepsilon)_l(-z-\varepsilon)_l}\,{}_4F_3\left[\begin{matrix}-l, & y+1+\varepsilon, & \alpha+1+y, & l+y-x-z \\ \alpha+1+y+\varepsilon, & 1+y-z, & 1+y-x & \end{matrix}\right]. $$

Note that in the above derivation we have applied the formula (3.10) twice. The kernel in (3.7) then takes the final form

(3.11)
$$ K(x,y,z) = A\frac{(\alpha+1+y+\varepsilon)_{x+z-y}}{(\alpha+1+y-z)_{x+z-y}}\sum_{l=0}^{x+z-y}\frac{(-x-z+y)_l(-N+y)_l}{l!(-\alpha-x-z-\varepsilon)_l} $$
$$ \cdot\,{}_4F_3\left[\begin{matrix}-l, & l+y-x-z, & y+1+\varepsilon, & \alpha+1+y \\ \alpha+1+y+\varepsilon, & 1+y-z, & 1+y-x & \end{matrix}\right]. $$

The series on the right hand side is evidently positive if $\varepsilon > -1$ and $\alpha \leqq -N-1-\varepsilon = \beta$. We have thus proved the following theorem:

THEOREM 1. *Let x, y, z be any three nonnegative integers between* 0 *and* N *satisfying the inequalities*

(3.12)
$$ 0 \leqq z \leqq x \leqq y \leqq x+z \leqq N. $$

Then the coefficients $E(x,y,z;\alpha,\beta,N)$ in (1.4) *are positive in the Eberlein region*

(3.13)
$$ \alpha \leqq \beta < -N. $$

Symmetry of $K(x,y,z)$, as is obvious in (1.8), and the positivity of the weight factor $\rho(z;\alpha,\beta,N)$ in (1.7) imply that the theorem also holds if any pair of the variables x, y, z in (3.12) is interchanged. Symmetry also implies that $K(x,y,z)=0$ if any one of the variables exceeds the sum of the other two.

4. Nonnegativity of $K(x, y, z)$ when $x+z>N$. As in case (i) we first reverse the ${}_4F_3$ series on the right of (2.21):

$$
\begin{aligned}
&{}_4F_3\left[\begin{matrix}-x-z+N-p, & -y-z+N-p, & N-z+1, & N-z+\beta+1\\ N-z+\alpha+1-p, & N-z+1-p, & -z\end{matrix}\right]\\
&=\frac{(-y-z+N-p)_{x+z+p-N}(N-z+1)_{x+z+p-N}(N-z+\beta+1)_{p+x+z-N}}{(N-z+\alpha+1-p)_{x+z+p-N}(N-z+1-p)_{x+z+p-N}(-z)_{x+z+p-N}}(-1)^{x+z+p-N}\\
&\quad\cdot{}_4F_3\left[\begin{matrix}-x-z+N-p, & -x, & N-x+1-p, & -\alpha-x\\ -x-p, & 1+y-x, & -\beta-x-p\end{matrix}\right].
\end{aligned}\tag{4.1}
$$

Use of (1.6) once again gives

$$
\begin{aligned}
&{}_4F_3\left[\begin{matrix}-x-z+N-p, & -x, & N-x-1+p, & -\alpha-x\\ -x-p, & 1+y-x, & -\beta-x-p\end{matrix}\right]\\
&=\sum_{k=0}^{x+z+p-N}\binom{x+z+p-N}{k}\frac{(-x)_k(-p)_{x+z+p-N-k}}{(-x-p)_{x+z-N+p}}\\
&\quad\cdot{}_3F_2\left[\begin{matrix}-k, & N-x+1-p, & -\alpha-x\\ 1+y-x, & -\beta-x-p\end{matrix}\right].
\end{aligned}
$$

The factor $(-p)_{x+z+p-N-k}$ vanishes unless $k\geqq x+z-N$. The series then actually starts from $k=x+z-N$. Making the transformation $k-(x+z-N)\to k$ we get

$$
\begin{aligned}
&{}_4F_3\left[\begin{matrix}-x-z+N-p, & -x, & N-x+1-p, & -\alpha-x\\ -x-p, & 1+y-x, & -\beta-x-p\end{matrix}\right]\\
&=\frac{(-x)_{x+z-N}(1+x+z-N)_p}{(-x-p)_{x+z-N+p}}(-1)^p\sum_{k=0}^{p}\frac{(-p)_k(-N+z)_k}{k!(1+x+z-N)_k}\\
&\quad\cdot{}_3F_2\left[\begin{matrix}-x-z+N-k, & N-x+1-p, & -\alpha-x\\ 1+y-x, & -\beta-x-p\end{matrix}\right]
\end{aligned}
$$

$$
\begin{aligned}
&=\frac{(-x)_{x+z-N}(1+x+z-N)_p}{(-x-p)_{x+z-N+p}}(-1)^p\sum_{k=0}^{p}\frac{(-p)_k(-N+z)_k(\alpha+1+y)_{x+z-N+k}}{k!(1+x+z-N)_k(1+y-x)_{x+z-N+k}}\\
&\quad\cdot{}_3F_2\left[\begin{matrix}-x-z+N-k, & -\beta-N-1, & -\alpha-x\\ -\beta-x-p, & -\alpha-x-y-z+N-k\end{matrix}\right]\\
&=\frac{(1+x+z-N)_p(\alpha+1+y)_{x+z-N}}{(1+x)_p(1+y-x)_{x+z-N}}\sum_{k=0}^{p}\frac{(-p)_k(-N+z)_k(\alpha+1+x+y+z-N)_k}{k!(1+x+z-N)_k(1+y+z-N)_k}\\
&\quad\cdot\sum_{l=0}^{x+z+k-N}\frac{(-x-z-k+N)_l(\varepsilon)_l(-\alpha-x)_l}{(N-x+1-p+\varepsilon)_l(-\alpha-x-y-z+N-k)_l l!}.
\end{aligned}\tag{4.2}
$$

The reduction in the second last step needed the use of (3.6). In the last line we have set $\beta=-N-1-\varepsilon$.

Setting (4.1) and (4.2) in (2.21) and carrying out some simplifications we obtain

$$
\begin{aligned}
K(x,y,z)&=B\frac{\Gamma(x+1+\varepsilon)\Gamma(N-x+1)}{\Gamma(N-x+1+\varepsilon)\Gamma(z+1)}\cdot\frac{(\alpha+1+y)_{x+z-N}}{(\alpha+1+N-z)_{x+z-N}}\\
&\quad\cdot\sum_{p=0}^{N-y}\frac{(-N+y)_p(-N+x-\varepsilon)_p}{(-N-\varepsilon)_p p!}\sum_{k=0}^{p}\frac{(-p)_k(-N+z)_k(\alpha+1+x+y+z-N)_k}{k!(1+x+z-N)_k(1+y+z-N)_k}\\
&\quad\cdot\sum_{l=0}^{x+z+k-N}\frac{(-x-z-k+N)_l(\varepsilon)_l(-\alpha-x)_l}{(N-x+1+\varepsilon-p)_l(-\alpha-x-y-z+N-k)_l l!}.
\end{aligned}\tag{4.3}
$$

The triple series on the right can be transformed as

$$\sum_{k=0}^{N-y}\sum_{l=0}^{x+z+k-N}\frac{(-N+z)_k(\alpha+1+x+y+z-N)_k(-x-z-k+N)_l(\varepsilon)_l(-\alpha-x)_l}{k!l!(1+x+z-N)_k(1+y+z-N)_k(-\alpha-x-y-z+N-k)_l}S_{k,l}$$

where

$$\begin{aligned}S_{k,l}&=\sum_{p=0}^{N-y-k}\frac{(-N+y)_{p+k}(-N+x-\varepsilon)_{p+k}(-p-k)_k}{(p+k)!(-N-\varepsilon)_{p+k}(N-x+1+\varepsilon-p-k)_l}\\&=\frac{(-N+y)_k(-N+x-\varepsilon)_{k-l}}{(-N-\varepsilon)_k}(-1)^{k+l}\sum_{p=0}^{N-y-k}\frac{(-N+y+k)_p(-N+x-\varepsilon+k-l)_p}{(-N-\varepsilon+k)_p p!}\\&=\frac{(-N+y)_k(-N+x-\varepsilon)_{k-l}(-x+l)_{N-y-k}}{(-N-\varepsilon)_k(-N-\varepsilon+k)_{N-y-k}}(-1)^{k+l}\\&=\frac{(-N+y)_k(-N+x-\varepsilon)_{k-l}(-x+l)_{N-y-k}}{(-N-\varepsilon)_{N-y}}(-1)^{k+l}\end{aligned}\tag{4.4}$$

by the Chu–Vandermonde theorem (2.9).

Using (4.4) in (4.3) and simplifying, we finally obtain

$$\begin{aligned}K(x,y,z)=C_\varepsilon\sum_{k=0}^{N-y}&\frac{(-N+x-\varepsilon)_k(-N+y)_k(-N+z)_k(\alpha+1+x+y+z-N)_k}{(1+x+y-N)_k(1+x+z-N)_k(1+y+z-N)_k k!}\\&\cdot{}_4F_3\left[\begin{matrix}-x-z-k+N, & \varepsilon, & -\alpha-x, & N-x-y-k\\ N-x+1+\varepsilon-k, & -x, & -\alpha-x-y-z+N-k\end{matrix}\right],\end{aligned}\tag{4.5}$$

where

$$C_\varepsilon=\left(\frac{x!}{N!}\right)^2\frac{y!z!(N-x)!\Gamma(y+1+\varepsilon)\Gamma(z+1+\varepsilon)(\alpha+1+y)_{x+z-N}\,(\alpha+1-N-\varepsilon)_N}{(x+z-N)!(z+y-N)!(y+x-N)!\Gamma(N-x+1+\varepsilon)\Gamma(N+\varepsilon+1)\,(\alpha+1+N-z)_{x+z-N}(\alpha+1)_z(\alpha+1)_{N-z}}.\tag{4.6}$$

The special case of $\varepsilon=0$ gives

$$\begin{aligned}K(x,y,z)&\equiv K(x,y,z;\alpha,-N-1,N)\\&=C_0\;{}_4F_3\left[\begin{matrix}x-N, & y-N, & z-N, & \alpha+1+x+y+z-N\\ 1+x+y-N, & 1+y+z-N, & 1+z+x-N\end{matrix}\right],\end{aligned}\tag{4.7}$$

where

$$C_0=\frac{\binom{x}{N-z}\binom{y}{N-x}\binom{z}{N-y}}{\binom{N}{x}\binom{N}{y}\binom{N}{z}}\cdot\frac{(\alpha+1-N)_N}{(\alpha+1)_z(\alpha+1)_{N-z}}\cdot\frac{(\alpha+1+y)_{x+z-N}}{(\alpha+1+N-z)_{x+z-N}}.\tag{4.8}$$

If $\alpha=-V+N-1$, V a positive integer $\geqq 2N$ as is the case in coding theory application [12], then K vanishes if $x+y+z>V$ and is evidently positive if $x+y+z\leqq V$. The kernel in (4.7) is essentially the same as p_{ijk} in Sloane [27, p. 244].

For a nonintegral $\alpha<-N-1$ the nonnegativity of K is not at all obvious in (4.7). This is because while $\alpha+1+x+y+z-N$ itself is negative when $x+y+z\leqq 2N$, the

Pochhammer product $(\alpha+1+x+y+z-N)_k$ may start having positive factors for some k, where $N-[(x+y+z)/2]<k\leqq N-y$.

Since the ${}_4F_3$ series in (4.5) is balanced, we can use (3.10) once again to transform it to another balanced ${}_4F_3$. There are, of course, many possibilities. But the form we finally wish to achieve is obtained by applying (3.10) in two stages. First, we employ the correspondence $m\leftrightarrow x+z+k-N$, $\xi\leftrightarrow N-x-y-k$, $\eta\leftrightarrow\varepsilon$, $\zeta\leftrightarrow-\alpha-x$, $u\leftrightarrow N-x+1+\varepsilon-k$, $v\leftrightarrow -x$ and $w\leftrightarrow N-\alpha-x-y-z-k$. In the second stage we use the correspondence $m\leftrightarrow x+z+k-N$, $\xi\leftrightarrow-\alpha-x$, $\eta\leftrightarrow N-x+1-k$, $\zeta\leftrightarrow 1+y+\varepsilon$, $u\leftrightarrow 1+y-x$, $v\leftrightarrow N-x+1+\varepsilon-k$, $w\leftrightarrow 1-\alpha-x-z+N-k$. Simplifying the resulting Pochhammer products we finally obtain

$$K(x,y,z)=D_\varepsilon\sum_{k=0}^{N-y}\frac{(y-N)_k(\alpha-\beta+x+y+z-2N)_k}{k!(1+x+z-N)_k}\cdot{}_4F_3\left[\begin{matrix}1+\alpha+y, & 1+y+\varepsilon, & k+y-N, & N-x-z-k\\ \alpha-\beta-N+y, & 1+y-x, & 1+y-z\end{matrix}\right] \tag{4.9}$$

where

$$D_\varepsilon=\frac{\binom{x}{N-y}\binom{y}{N-x}\binom{z}{N-y}}{\binom{N}{x}\binom{N}{y}\binom{N}{z}}\frac{(y-z)!(y-x)!N!}{y!(x+y-N)!(y+z-N)!}\cdot\frac{\Gamma(y+1+\varepsilon)(\alpha+1-N-\varepsilon)_N}{\Gamma(N+1+\varepsilon)(\alpha+1)_z(\alpha+1)_{N-z}}\frac{(\alpha-\beta-N+y)_{x+z-N}}{(\alpha+1+N-z)_{x+z-N}}. \tag{4.10}$$

In (4.9) and (4.10) we have exhibited the explicit dependence on $\alpha-\beta$ at places where it plays a crucial role.

It can be seen from the above expressions that a necessary condition when α, $\beta<-N$ for the nonnegativity of $K(x,y,z)$ is that $\alpha\leqq\beta<-N$. For, suppose $y=N$, $x+z=N+1$. Then, from (4.9) and (4.10) we obtain

$$K(x,N,N+1-x)=\left[x^2\binom{N}{x}\right]^{-1}\frac{(\alpha+1-N-\varepsilon)_N}{(\alpha+1)_{N+1-x}(\alpha+1)_{x-1}}\frac{\alpha-\beta}{\alpha+x},\qquad x>1.$$

This is negative if $\beta<\alpha<-N$. For a nonnegative representation of $K(x,y,z)$ it is therefore necessary to restrict to the region $\alpha\leqq\beta<-N$.

Apart from the factor D_ε in (4.9) the expression that seems to control the sign pattern of the terms of the finite series is $(\alpha-\beta+x+y+z-2N)_k$. Accordingly, we shall distinguish between the two cases: (i) $x+y+z\leqq 2N$ and (ii) $2N<x+y+z\leqq 3N$.

Case (i). $x+y+z\leqq 2N$. For $\alpha\leqq\beta<-N$ the ${}_4F_3$ function in (4.9) is obviously positive under the restrictions on x, y, z stated before. Also, if $\beta-\alpha$ is a nonnegative integer the series for $K(x,y,z)$ is a series of nonnegative terms and hence the kernel K is positive since D_ε is. However, the situation is not so straightforward if $\beta-\alpha$ is not an integer. First, let us split the series in (4.9) into two parts:

$$\sum_{k=0}^{N-y}=\sum_{k=0}^{M}+\sum_{k=M+1}^{N-y}$$

where $M\equiv N-[(x+y+z)/2]$. The first series on the right is obviously positive. The terms in the second series are also positive in the range $\beta-\alpha\geqq N-1$, $\beta<-N$. We need, therefore, to consider the second series only in the region $0<\beta-\alpha<N-1$.

Even in this region the product $(\alpha-\beta+x+y+z-2N)_k$ has positive factors only when $k \geqq 2N-x-y-z+\beta-\alpha+1$. Thus the nonnegativity of $K(x, y, z)$ hinges on the sign of the series

$$\sum_{k=2M+[b]}^{N-y} \frac{(y-N)_k(\alpha-\beta+x+y+z-2N)_k}{k!(1+x+z-N)_k} \cdot {}_4F_3\left[\begin{matrix}1+\alpha+y, & 1+y+\varepsilon, & N-x-z-k, & k+y-N\\ \alpha-\beta-N+y, & 1+y-x, & 1+y-z & \end{matrix}\right]$$

$$= \frac{(y-N)_{2M+[b]}(\alpha-\beta+x+y+z-2N)_{2M+[b]}}{(2M+[b])!(1+x+z-N)_{2M+[b]}}$$

$$(4.11)\quad \sum_{l=0}^{N-y-2M-[b]} \frac{(y-N+2M+[b])_l([b]-b+2\mu)_l}{(2M+[b]+1)_l(1+x+z-N+2M+[b])_l} \cdot {}_4F_3\left[\begin{matrix}1+\alpha+y, & 1+y+\varepsilon, & N-2M-[b]-x-z-l, & l+2M-N+[b]+y\\ \alpha-\beta-N+y, & 1+y-x, & 1+y-z & \end{matrix}\right],$$

where $b=\beta-\alpha$ and $\mu=0$ or $\frac{1}{2}$ according as $x+y+z$ is even or odd.

While the factor in front of the series is positive the terms of the series itself have alternating sign with the first term positive if $\mu=\frac{1}{2}$ and the first two terms positive if $\mu=0$. However, if we denote the ${}_4F_3$ function by a_l then

$$a_l-a_{l+1}=2\sum_{m=1}^{N-y-2M-[b]-l} m\frac{(1+\alpha+y)_m(1+y+\varepsilon)_m(1+2M-N+[b]+y+l)_{m-1}(N-2M-[b]-x-z-l)_{m-1}}{m!(\alpha-\beta-N+y)_m(1+y-x)_m(1+y-z)_m} \cdot\left\{l+2M+[b]+\frac{x+y+z+1}{2}-N\right\},$$

$l \geqq 0$. Since the expression within the curly brackets is positive the series on the right has only nonnegative terms. Therefore, $a_l \geqq a_{l+1}$, $l \geqq 0$. Also,

$$0 \leqq N-y-2M-[b] < 1+x+z-N+2M+[b] \quad \text{and} \quad 2\mu+[b]-b<1+2M+[b].$$

Hence the terms in the alternating series above are monotonic decreasing. The series on the right of (4.11) is, therefore, positive.

Case (ii). $2N < x+y+z \leqq 3N$. If $\varepsilon>-1$ and $\beta-\alpha$ is a nonnegative integer, then, because of the factor $(\alpha-\beta-N+y)_{x+z-N}$ in D_ε, the kernel $K(x, y, z)$ vanishes when $x+y+z-2N \geqq \beta-\alpha+1$. On the other hand, if $\beta-\alpha$ is still a nonnegative integer but $0<x+y+z-2N \leqq \beta-\alpha$ then the terms on the right of (4.9) are nonnegative and hence K is positive.

However, the situation changes radically in $\beta-\alpha$ is not a nonnegative integer. In this case $K(x, y, z)$ remains positive for all permissible values of x, y, z if $\beta-\alpha> N-1$, $\beta<-N$, as can be verified directly from (4.9) and (4.10). For $0<\beta-\alpha<N-1$ let us consider the example: $x=y=z=N$, $\beta=-N-1$. Then from (4.7) we obtain

$$K(N, N, N; \alpha, -N-1, N)=\frac{(\alpha+1-N)_N}{(\alpha+1)_N}\cdot\frac{(\alpha-\beta)_N}{(\alpha+1)_N}.$$

This can be either positive or negative.

We summarize the results of this section in the following theorem:

THEOREM 2. *Let* x, y, z *be any three nonnegative integers satisfying the inequalities*

$$0 \leqq z \leqq x \leqq y \leqq N, \qquad x+z>N. \tag{4.12}$$

Then:

(i) *If* $\beta<\alpha<-N$, *some of the coefficients* $E(x, y, z; \alpha, \beta, N)$ *are negative.*

(ii) *If* $\alpha \leqq \beta<-N$ *and* $\beta-\alpha$ *is a nonnegative integer then* E *is nonnegative. In particular,* E *vanishes if* $x+y+z-2N \geqq \beta-\alpha+1$.

(iii) *If* $\beta-\alpha$ *is not an integer and* $\alpha<\beta<-N$ *then* E *is positive for all values of* x, y, z *satisfying* (4.12) *provided* $\beta-\alpha>N-1$. *When* $0<\beta-\alpha<N-1$, E *is positive in the region* $x+y+z \leqq 2N$ *but can have any sign when* $2N<x+y+z \leqq 3N$.

Note. At the time of writing this paper the author was unaware of Professor Charles F. Dunkl's recent work containing a double series representation of $K(x, y, z)$. His results, obtained by group-theoretic methods, agree with ours in the special case $\beta=-N-1$, but for $\beta+N+1 \leqq -1$ his double series is not quite the same as ours. His double series is also expressible in terms of ${}_4F_3$ functions, but unlike our results, they are not balanced.

Professor Dunkl's paper, entitled *Spherical functions on compact groups and applications to special functions*, is to appear in the Proceedings of Rome Conference on Harmonic Analysis, March 1976.

Acknowledgment. I would like to thank Professors R. Askey and G. Gasper for bringing Dunkl's paper to my attention. Thanks are also due to Professor Dunkl, who sent over a preprint of his paper as well as some valuable comments.

REFERENCES

[1] R. ASKEY, *Orthogonal expansions with positive coefficients*, Proc. Amer. Math. Soc., 16 (1965), pp. 1191–1194.

[2] ———, *Jacobi polynomial expansions with positive coefficients and imbeddings of projective spaces*, Bull. Amer. Math. Soc., 74 (1968), pp. 301–304.

[3] ———, *Orthogonal polynomials and positivity*, Special Functions and Wave Propagation, Studies in Applied Mathematics, 6, D. Ludwig and F. W. J. Olver, eds., Society for Industrial and Applied Mathematics, Philadelphia, 1970, pp. 64–85.

[4] ———, *Orthogonal expansions with positive coefficients. II*, this Journal, 2 (1971), pp. 340–346.

[5] ———, *Orthogonal Polynomials and Special Functions*, Regional Conference Series in Applied Mathematics, vol. 21, Society for Industrial and Applied Mathematics, Philadelphia, 1975.

[6] R. ASKEY AND G. GASPER, *Jacobi polynomial expansions of Jacobi polynomials with non-negative coefficients*, Proc. Cambridge Philos. Soc., 70 (1971), pp. 243–255.

[7] R. ASKEY, G. GASPER AND M. E. H. ISMAIL, *A positive sum from summability theory*, J. Approximation Theory, 13 (1975), pp. 413–420.

[8] R. ASKEY AND J. STEINIG, *Some positive trigonometric sums*, Trans. Amer. Math. Soc., 187 (1974), pp. 295–307.

[9] R. ASKEY AND S. WAINGER, *A convolution structure for Jacobi series*, Amer. J. Math., 91 (1969), pp. 463–485.

[10] W. N. BAILEY, *Generalized Hypergeometric Series*, Stechert–Hafner Service Agency, New York and London, 1964.

[11] R. D. COOPER, M. R. HOARE AND M. RAHMAN, *Stochastic processes and special functions: On the probabilistic origin of some positive kernels associated with classical orthogonal polynomials*, J. Math. Anal. Appl., 61 (1977), pp. 262–291.

[12] P. DELSARTE, *An algebraic approach to the association schemes of coding theory*, Philips Research Reports Supplements, no. 10, 1973.

[13] P. J. EBERLEIN, *A two parameter test matrix*, Math. Comput., 18 (1964), pp. 296–298.

[14] G. GASPER, *Linearization of the product of Jacobi polynomials II*, Canad. J. Math., 22 (1970), pp. 582–593.
[15] ———, *Positivity and the convolution structure for Jacobi series*, Ann. of Math., 93 (1971), pp. 112–118.
[16] ———, *Banach algebras for Jacobi series and positivity of a kernel*, Ibid., 95 (1972), pp. 261–280.
[17] ———, *Non-negativity of a discrete Poisson kernel for the Hahn polynomials*, J. Math. Anal. Appl., 42 (1973), pp. 438–451.
[18] ———, *Projection formulas for orthogonal polynomials of a discrete variable*, Ibid., 45 (1974), pp. 176–198.
[19] ———, *Positive integrals of Bessel functions*, this Journal, 5 (1975), pp. 868–881.
[20] ———, *Positivity and special functions*, Theory and Application of Special Functions, R. Askey, ed., Academic Press, New York, 1975, pp. 375–433.
[21] P. HARTMAN AND G. WATSON, *"Normal" distribution functions on spheres and the modified Bessel functions*, Ann. Probability, 2 (1974), pp. 593–607.
[22] M. RAHMAN, *Construction of a family of positive kernels from Jacobi polynomials*, this Journal, 7 (1976), pp. 92–116.
[23] ———, *A five-parameter family of positive kernels from Jacobi polynomials*, this Journal, 7 (1976), pp. 386–413.
[24] ———, *Some positive kernels and bilinear sums for Hahn polynomials*, this Journal, 7 (1976), pp. 414–435.
[25] ———, *On a generalization of the Poisson kernel for Jacobi polynomials*, this Journal, 8 (1977), pp. 1014–1031.
[26] L. J. SLATER, *Generalized Hypergeometric Functions*, Cambridge University Press, Cambridge, England, 1966.
[27] N. J. A. SLOANE, *An introduction to association schemes and coding theory*, Theory and Application of Special Functions, R. Askey, ed., Academic Press, New York, 1975, pp. 225–260.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0014 $01.00/0

INTERPOLATION AND APPROXIMATION OF GENERALIZED AXISYMMETRIC POTENTIALS*

ALLAN J. FRYANT†

Abstract. Using a function theoretic approach, this paper deals with the uniform approximation of generalized axisymmetric potentials (GASP's) by GASP polynomials. Analogues of the Runge and Walsh polynomial approximation theorems are obtained. Approximation of a GASP which is regular on the closure of a region by GASP polynomials interpolating on the boundary is considered, where conditions on the choice of interpolation points which yield the interpolants uniformly convergent at a geometric rate are found. Solution of the Dirichlet problem by GASP polynomial interpolation to the boundary values is discussed, and a constructive solution given in the case of an ellipse.

1. Introduction. Solutions of the equation

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{2\mu}{y}\frac{\partial u}{\partial y}=0, \tag{1.1}$$

where μ is a nonnegative real number, are called *generalized axisymmetric potentials* (GASP's). In the case where $2\mu = n-2$, $n=2, 3, \cdots$, these functions are harmonic in R^n, i.e. satisfy Laplace's equation

$$\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}+\cdots+\frac{\partial^2 u}{\partial x_n^2}=0,$$

with $x=x_1$ and $y=(x_2^2+\cdots+x_n^2)^{1/2}$. The case of noninteger values of 2μ was first extensively studied by Weinstein, and was motivated by a number of applications (see [22] and references included there). For example the well-known Tricomi equation can be transformed into (1.1) with $\mu=\frac{1}{6}$.

Let Ω be a region in the x, y-plane which is symmetric with respect to the x-axis. We say that u is a GASP which is *regular* in Ω if $u\in C^2(\Omega)$, satisfies (1.1) for all $(x, y)\in\Omega$, $y\neq 0$, and $\partial u(x, 0)/\partial y=0$ on the intersection of Ω with the x-axis. Such functions are necessarily symmetric, i.e. satisfy $u(x, y)=u(x, -y)$. We say that u is regular on cl (Ω) if u is regular in some region $\Omega'\supset \text{cl}(\Omega)$.

In developing approximation theoretic results for GASP's, Gilbert's A_μ integral operator [10, pp. 165–169] will be used as a principle tool. The A_μ integral operator transforms analytic functions of a single complex variable to GASP's:

$$u(x, y)=A_\mu(f)=\alpha_\mu\int_L f(z)(\zeta-\zeta^{-1})^{2\mu-1}\zeta^{-1}\,d\zeta, \tag{1.2}$$

where $z=x+iy(\zeta+\zeta^{-1})/2$, $L=\{e^{i\phi}: 0\leqq\phi\leqq\pi\}$, and $\alpha_\mu=[\int_L(\zeta-\zeta^{-1})^{2\mu-1}\zeta^{-1}\,d\zeta]^{-1}$.

If u is a GASP which is regular in a neighborhood of the origin, it has the "ultra-spherical harmonic" expansion

$$u(x, y)=\sum_{n=0}^{\infty}\frac{\Gamma(2\mu)\Gamma(n+1)}{\Gamma(n+2\mu)}a_n r^n C_n^\mu(\cos\theta),$$

where $(x, y)=(r\cos\theta, r\sin\theta)$, and C_n^μ are the Gegenbauer polynomials. The corresponding A_μ associate is then $f(z)=\sum_{n=0}^\infty a_n z^n$. It should be noted that the A_μ

* Received by the editors September 20, 1976.

† Mathematics Department, U.S. Naval Academy, Annapolis, Maryland 21402. This work was supported by a grant from the U.S. Naval Academy Research Council.

associate of a GASP is the analytic continuation of its restriction to the axis of symmetry, i.e. $f(z) = u(z, 0)$.

The operation of "taking the real part" provides a transformation from analytic functions of a single complex variable to harmonic functions in two dimensions, and thus allows the development of an approximation theory for harmonic functions which is based upon results from constructive function theory (see [18]–[20]). In a similar manner the A_μ integral operator will be used to transform results regarding the polynomial approximation of analytic functions to corresponding approximation theoretic results for GASP's. Uniform convergence will be considered throughout the paper, where for a given region Ω, $\|\cdot\|$ will always denote the uniform norm over $\mathrm{cl}(\Omega)$. A region Ω will be called *axiconvex* if $x+iy \in \Omega$ implies $x+i\lambda y \in \Omega$ for every $\lambda \in [-1, 1]$.

2. GASP polynomial approximation. For convenience we express $r^n C_n^\mu(\cos\theta)$ in rectangular coordinates, i.e. define $C_n^\mu(x, y) = r^n C_n^\mu(\cos\theta)$. $C_n^\mu(x, y)$ is a homogeneous polynomial of degree n in x, y and satisfies equation (1.1). A GASP polynomial of degree n is then a sum of the form $\sum_{k=0}^n a_k C_k^\mu(x, y)$, where $a_n \neq 0$.

Using the A_μ integral operator, we first obtain a Runge theorem for GASP's as an immediate consequence Runge's function theoretic result.

THEOREM 2.1. *Let Ω be a bounded axiconvex region and u be a GASP which is regular on* $\mathrm{cl}(\Omega)$. *Then there exist GASP polynomials q_n such that $q_n \to u$ uniformly on* $\mathrm{cl}(\Omega)$.

Proof. Gilbert has shown [10, pp. 177–179] that a point (x, y) is a singularity of a GASP $u(x, y)$ if and only if $x \pm iy$ is a singularity of its A_μ associate. Thus the axiconvex regions of regularity of u and its A_μ associate f coincide (this later result has also been obtained by Erdélyi [8]). Thus f is analytic on $\mathrm{cl}(\Omega)$, and by Runge's theorem there exist polynomials p_n which converge uniformly to f on $\mathrm{cl}(\Omega)$. Let $q_n = A_\mu(p_n)$. Then q_n is a GASP polynomial (of the same degree as p_n), and for every $(x, y) \in \mathrm{cl}(\Omega)$

$$\begin{aligned}|u(x, y) - q_n(x, y)| &= |A_\mu(f - p_n)| \\ &= \left|(-1)^\mu 2^{2\mu-1}\alpha_\mu \int_0^\pi [f(x+iy\cos t) - p_n(x+iy\cos t)](\sin t)^{2\mu-1}\,dt\right| \\ &\leqq \|f - p_n\|.\end{aligned}$$

Therefore $\|u - q_n\| \leqq \|f - p_n\| \to 0$.

We next show that the approximating polynomials can be chosen so as to converge at a geometric rate, and determine the maximal degree of this convergence.

THEOREM 2.2. *Let Ω be a bounded axiconvex region, and ϕ be the conformal mapping of Ω onto the exterior of the unit disk, with $\phi(\infty) = \infty$. Let u be a GASP which is regular on* $\mathrm{cl}(\Omega)$. *If ρ is the largest number for which u is regular interior to $\Gamma_\rho = \{z : |\phi(z)| = \rho\}$, then for every R where $1 < R < \rho$ there exist GASP polynomials q_n of degree n such that*

$$\|u - q_n\| \leqq M/R^n, \tag{2.1}$$

where M is a constant depending only on R. Further this degree of convergence need not obtain for $R > \rho$.

Proof. Since u is regular on $\mathrm{cl}(\Omega)$ it follows that $\rho > 1$. Further, since u is regular interior to Γ_ρ, its A_μ associate is also analytic on this region. Thus for $1 < R < \rho$ there exist polynomials p_n of degree n such that $\|f - p_n\| \leqq M/R^n$, where M is a constant depending only on R [20, p. 79]. Defining $q_n = A_\mu(p_n)$ and arguing on the continuity of the linear operator A_μ as in the preceding theorem then yields (2.1).

To show that this degree of convergence need not obtain for $R > \rho$, consider the case where $\Omega = D_r$, an open disk of radius r centered at the origin. Then $\Gamma_\rho = \{z : |z| = r\rho\}$, and since the axiconvex regions of regularity of u and its A_μ associate f coincide, Γ_ρ is the circle of least radius which is centered at the origin and passes through a singularity f.

Suppose that for some $R > \rho$ there exist GASP polynomials q_n of degree n satisfying (2.1) where $\|\cdot\|$ is the uniform norm over cl (D_r). Let $p_n = A_\mu^{-1}(q_n)$, and let $r' < r$ be the chosen sufficiently large so that $Rr' > r\rho$. Then for $z \in D_{r'}$,

$$f(z) - p_n(z) = A_\mu^{-1}(u - q_n) = \int_{-1}^{1} [u(r, \xi) - q_n(r, \xi)] K(z/r, \xi)\, d\xi,$$

where

$$K(z/r, \xi) = \frac{\mu \Gamma(\mu)^2 2^{2\mu-1}(1-\xi^2)^{\mu-1/2}(1-z^2/r^2)}{\pi[1-2\xi(z/r)+z^2/r^2]^{\mu+1}}.$$

(For a development of this integral representation of the inverse transform A_μ^{-1}, see [10, p. 173].) Now $K(z/r, \xi)$ is uniformly bounded for all $z \in D_{r'}$, $\xi \in [-1, 1]$. Thus

$$\|f - p_n\| \leqq k\|u - q_n\| \leqq kM/R^n,$$

where k is a constant independent of n. But since p_n is a polynomial of degree n, this degree of convergence implies that f can be analytically continued throughout the disk $|z| < Rr'$ (see [17, p. 21]). Since $Rr' > r\rho$ and f has a singularity on the circle $|z| = r\rho$, we have a contradiction.

Using the A_μ integral operator, a great number of results regarding various types of polynomial expansions of analytic functions can also be transformed to analogous results regarding GASP's. For example the A_μ transforms of the Faber polynomials yield GASP polynomials which retain the usual expansion properties associated with the Faber polynomials. Consider a bounded, simply connected region Ω, where $w = \phi(z) = z + a_0 + a_1 z^{-1} + \cdots$ maps the compliment of Ω onto the circular region $|w| > \rho$, and let $\psi = \phi^{-1}$. The Faber polynomials $f_n(z)$ associated with the region Ω are defined by

$$\frac{\psi'(t)}{\psi(t) - z} = \sum_{n=0}^{\infty} f_n(z) t^{-n}, \qquad [17, \text{p. } 130].$$

For a bounded axiconvex region Ω, we define the *GASP Faber polynomials* u_n as the A_μ transforms of the Faber polynomials for Ω:

$$u_n(x, y) = A_\mu(f_n) = \alpha_\mu \int_0^\pi f_n(x + iy \cos t)(\sin t)^{2\mu-1}\, dt.$$

Appealing to the well known expansion properties of the Faber polynomials [17, p. 138], we obtain the following result immediately.

THEOREM 2.3. *Let Ω be a bounded axiconvex region, and $w = \phi(z) = z + a_0 + a_1 z^{-1} + \cdots$ map the compliment of Ω onto the circular region $|w| > \rho$, and let $\psi = \phi^{-1}$. If u is a GASP which is regular on the region B_R bounded by $\Gamma_R = \{z : |\phi(z)| = R\}$, then u can be expanded into a series of GASP Faber polynomials*

$$u(x, y) = \sum_{n=0}^{\infty} a_n u_n(x, y) \tag{2.2}$$

where

$$a_n = \frac{1}{2\pi i}\int_{|z|=r} u(\psi(z),0)z^{-k-1}\,dz, \qquad \rho<r<R.$$

Further,

$$\limsup_{n\to\infty} |a_n|^{1/n} = 1/R,$$

and the series (2.2) *converges uniformly on compact subsets of* B_R.

We note that for a disk centered at the origin, the GASP Faber polynomials are $C_n^\mu(x, y) = r^n C_n^\mu(\cos\theta)$, and (2.2) becomes the usual "ultra-spherical harmonic" expansion of u on the disk.

A region will be called *star-shaped* if it is convex with respect to a point on the x-axis. Using the result of Theorem 2.1 and the fact that GASP's are invariant under homothetic transformation yields the following analogue of Walsh's polynomial approximation theorem.

THEOREM 2.4. *Let* Ω *be a bounded, axiconvex, star-shaped region, and* u *be a GASP which is regular in* Ω *and continuous on* $\mathrm{cl}(\Omega)$. *Then there exist GASP polynomials* q_n *such that* $q_n \to u$ *uniformly on* $\mathrm{cl}(\Omega)$.

Proof. Without loss of generality, assume Ω is star shaped with respect to the origin. Express u in polar coordinates $u(r, \theta)$. Since GASP's are invariant under homothetic transformation, the functions

$$u_n(r,\theta) = u\left(\frac{n}{n+1}r, \theta\right), \qquad n = 1, 2, \cdots,$$

are GASP's which are regular on $\mathrm{cl}(\Omega)$. Let $\varepsilon > 0$. Since u is uniformly continuous on $\mathrm{cl}(\Omega)$, there exists an n such that $\|u - u_n\| < \varepsilon/2$. Further, since u_n is regular on $\mathrm{cl}(\Omega)$, Theorem 2.1 provides a GASP polynomial q_k for which $\|u_n - q_k\| < \varepsilon/2$. Thus $\|u - q_k\| < \varepsilon$.

3. Approximation by GASP polynomial interpolation. We next consider the uniform approximation of a GASP which is regular on the closure of a region by GASP polynomials which interpolate it at points on the boundary.

For $C_n = \{(x_k, y_k)\}_{k=0}^n$ we define

$$V(C_n) = \det [C_k^\mu(x_j, y_j)], \qquad k, j = 0, 1, \cdots, n,$$

and

$$V_k(x, y; C_n) = V(C_n)|_{(x_k, y_k)=(x,y)}.$$

Let Ω be a region, u be any function defined on $\partial\Omega$, and $C_n \subset \partial\Omega$. If $V(C_n) \neq 0$, then clearly

$$L_n(x, y; C_n; u) = \sum_{k=0}^n u(x_k, y_k) V_k(x, y; C_n)/V(C_n)$$

is (for a given value of μ) the unique GASP polynomial of degree n which interpolates u at the points (x_k, y_k), $k = 0, 1, \cdots, n$.

We first consider the case of interpolation on a disk, obtaining an Hermite type formula for the error in approximation by GASP polynomial interpolants. For convenience results are stated for the unit disk, $D_1 = \{z: |z| < 1\}$. Note that if u is a GASP which is regular on $\mathrm{cl}(D_1)$, then u has the expansion $u(x, y) = \sum_{k=0}^\infty a_k C_k^\mu(x, y)$, where

$\limsup |a_k|^{1/k} = 1/R$, $R > 1$, and the series converges uniformly on compact subsets of $B_R = \{z: |z| < R\}$. Also if $C_n = \{(x_k^{(n)}, y_k^{(n)})\}_{k=0}^n \subset \partial D_1$, and $x_k^{(n)} \neq x_j^{(n)}$ for $k \neq j$, then $V(C_n) \neq 0$.

THEOREM 3.1. *Let $u(x, y) = \sum_{k=0}^\infty a_k C_k^\mu(x, y)$ be a GASP which is regular on $\mathrm{cl}(D_1)$, and $C_n = \{(x_k^{(n)}, y_k^{(n)})\}_{k=0}^n \subset \partial D_1$, where $x_k^{(n)} \neq x_j^{(n)}$ for $k \neq j$. Then on the unit circle $x^2 + y^2 = 1$,*

$$u(x, y) - L_n(x, y; C_n; u) = \frac{1}{2\pi i} \int_\gamma \frac{w_n(x) h(\zeta)}{w_n(\zeta)(\zeta - x)} d\zeta,$$

where

$$w_n(\zeta) = \prod_{k=0}^n (\zeta - x_k^{(n)}), \qquad h(\zeta) = u(\zeta, \sqrt{1-\zeta^2}) = \sum_{k=0}^\infty a_k C_k^\mu(\zeta),$$

and γ is any simple closed rectifiable curve about $[-1, 1]$ lying within the ellipse $E_R = \{(R e^{i\phi} + R^{-1} e^{-i\phi})/2 : 0 \leqq \phi \leqq 2\pi\}$, where $\limsup |a_n|^{1/n} = 1/R$.

Proof. Since $C_n^\mu(x, y) = C_n^\mu(x/r)$, $r^2 = x^2 + y^2$, we have $C_n^\mu(x, \sqrt{1-x^2}) = C_n^\mu(x)$, and thus

$$u(x, \sqrt{1-x^2}) = \sum_{n=0}^\infty a_n C_n^\mu(x). \tag{3.3}$$

With the use of Bernstein's bound on the growth of polynomials in the complex plane [11, p. 42], the well known inequality

$$|C_n^\mu(x)| \leqq \frac{\Gamma(n+2\mu)}{\Gamma(n+1)\Gamma(2\mu)}, \qquad x \in [-1, 1],$$

implies

$$|C_n^\mu(\zeta)| \leqq \frac{\Gamma(n+2\mu)}{\Gamma(n+1)\Gamma(2\mu)} R^n, \qquad \zeta \in E_R.$$

Since

$$\lim_{n\to\infty} \left[\frac{\Gamma(n+2\mu)}{\Gamma(n+1)\Gamma(2\mu)}\right]^{1/n} = 1,$$

this implies that the function given by (3.3) can be analytically continued throughout the region Ω_R interior to E_R, for the series $\sum_{n=0}^\infty a_n C_n^\mu(\zeta)$ converges uniformly for ζ restricted to compact subsets of Ω_R. Further, since $L_n(x, y; C_n; u)$ is a polynomial of degree n in x and y and is even in y, $L_n(z, \sqrt{1-z^2}; C_n; u)$ is a polynomial of degree n in z, and interpolates $h(z) = u(z, \sqrt{1-z^2})$ at the points $x_j^{(n)} \in [-1, 1]$, $j = 0, 1, \cdots, n$. Thus by the Hermite remainder formula [20, p. 50],

$$u(z, \sqrt{1-z^2}) - L_n(z, \sqrt{1-z^2}; C_n; u) = \frac{1}{2\pi i} \int_\gamma \frac{w_n(z) h(\zeta)}{w_n(\zeta)(\zeta - z)} d\zeta,$$

where γ is a curve lying within E_R and z is interior to γ. In particular if γ is a curve about $[-1, 1]$, we have on the circle $x^2 + y^2 = 1$,

$$u(x, y) - L_n(x, y; C_n; u) = \frac{1}{2\pi i} \int_\gamma \frac{w_n(x) h(\zeta)}{w_n(\zeta)(\zeta - x)} d\zeta.$$

Since a nonconstant GASP which is regular in a region cannot attain a maximum at an interior point [14], the result of this theorem provides a bound on the difference

$u-L_n$ throughout the disk. That is,

$$\|u-L_n\| \leqq \max_{x\in[-1,1]} \frac{1}{2\pi}\left|\int_\gamma \frac{w_n(x)h(\zeta)}{w_n(\zeta)(\zeta-x)}\,d\zeta\right|. \tag{3.1}$$

By using this bound it is possible to determine necessary and sufficient conditions on the choice of interpolation points $C_n \subset \partial D_1$ which yield GASP interpolating polynomials that converge uniformly on $\mathrm{cl}\,(D_1)$. The following two theorems were first obtained for axisymmetric harmonic functions in R^3 (that is, solutions of the equation (1.1) in the case $\mu=\frac{1}{2}$), [9]. Since the results for the general case $\mu>0$ follow from the bound (3.1) in an entirely similar manner, we omit their proofs here.

THEOREM 3.2. *Let* $C_n=\{(x_k^{(n)}, y_k^{(n)})\}_{k=0}^n \subset \partial D_1$, *where* $x_k^{(n)} \neq x_j^{(n)}$ *for* $k\neq j$. *Also let* $w_n(x)=\prod_{k=0}^n (x-x_k^{(n)})$, *and* $\theta_n(z)=2w_n(z)^{1/(n+1)}/\phi(z)$, *where* $\phi(z)=z+\sqrt{z^2-1}$ *and the branch is chosen so that* $\phi(\infty)=\infty$. *Then*

$$L_n(x, y; C_n; u)\to u(x, y)$$

uniformly on $\mathrm{cl}\,(D_1)$ *for every GASP u which is regular on* $\mathrm{cl}\,(D_1)$ *if and only if either*

$$\text{(i)} \qquad |\theta_n(z)|\to 1$$

uniformly on compact subsets of $C_\infty[-1, 1]$ (C_∞ = *extended complex plane*), *or*

$$\text{(ii)} \qquad \lim_{n\to\infty} M_n^{1/(n+1)}=1/2,$$

where $M_n=\max|w_n(x)|$, $x\in[-1,1]$. *Further, in this case* $\|u-L_n\|<M/R^n$, *where R is any number greater than* 1 *and less than the radius of the largest disk centered at the origin in which u is regular, and M is a constant depending only on R.*

THEOREM 3.3. *Let* $C_n=\{(x_k^{(n)}, y_k^{(n)})\}_{k=0}^n \subset \partial D_1$, *where* $x_k^{(n)}\neq x_j^{(n)}$ *for* $k\neq j$. *Then* $L_n\{x, y; c_n; u)\to u(x, y)$ *uniformly on* $\mathrm{cl}\,(B_1)$ *for all GASP's u which are regular on* $\mathrm{cl}\,(B_1)$ *if and only if for every* (x_0, y_0) *such that* $x_0^2+y_0^2>1$, *the GASP interpolating polynomials* $L_n(x, y; C_n, V_j)$, $j=1, 2$, *converge uniformly on* $\mathrm{cl}\,(B_1)$ *to the solutions of the respective Dirichlet problems having boundary values*

$$V_1(x, y)=\frac{1}{(x-x_0)^2+y_0^2}$$

and

$$V_2(x, y)=\frac{x}{(x-x_0)^2+y_0^2}, \qquad x^2+y^2=1.$$

Thus, for example, the interpolants converge uniformly at a geometric rate for the following choices of $x_k^{(n)}$, $k=0, 1, 2, \cdots, n$.

1) *Vandermonde points*: $x_k^{(n)}$ chosen to maximize the Vandermonde determinant

$$|v(x_0, x_2, \cdots, x_n)|=|\det[1, x_k, x_k^2, \cdots, x_k^n]_{k=0}^n|,$$

over the interval $[-1, 1]$.

2) *Fejer points*: $x_k^{(n)}=\phi(z_k)$, where $\phi(z)=(z+z^{-1})/2$ and z_k are uniformly distributed (see [17, pp. 22–28]) points on the unit circle.

3) *Chebyshev points*: $x_k^{(n)}=\cos[(2k+1)\pi/(2n+2)]$, the zeros of the Chebyshev polynomials.

Appealing to the result of Theorem 2.2, we next determine a condition on the choice of interpolation points which yields uniform convergence in the case of an arbitrary bounded axiconvex region.

THEOREM 3.4. *Let Ω be a bounded axiconvex region, and $C_n^* = \{(x_k^{(n)}, y_k^{(n)})\}_{k=0}^n$ maximize the determinant $|V(C_n)|$ over $\partial\Omega$. Then*

$$\lim_{n\to\infty} L_n(x, y; C_n^*; u) = u(x, y)$$

uniformly on cl (Ω) *for every GASP u which is regular on* cl (Ω). *Further, the convergence is at a geometric rate, i.e.*

$$\|u - L_n\| \leqq (n+2)M/R^n$$

where M and R are as given by Theorem 2.2.

Proof. First note that since Ω is a bounded axiconvex region, for every n, $V(C_n^*) \neq 0$. (This is easily proved by induction and appeal to the maximum principle for GASP's). Therefore given any $n+1$ values a_k, $\sum_{k=0}^n a_k V_k(x, y; C_n^*)/V(C_n^*)$ is the unique GASP of degree n which assumes these values at the points $(x_k^{(n)}, y_k^{(n)})$, $k = 0, 1, \cdots, n$. In particular, if $p_n(x, y)$ is a GASP of degree n,

$$L_n(x, y; C_n^*; p_n) \equiv p_n(x, y).$$

Let q_n be the GASP polynomial approximants to u given by Theorem 2.2. Then for $(x, y) \in$ cl (Ω),

$$\begin{aligned} &|u(x, y) - L_n(x, y; C_n^*, U)| \\ &= \left| \sum_{k=0}^n u(x_k, y_k) \frac{V_k(x, y; C_n^*)}{V(C_n^*)} - \sum_{k=0}^n q_n(x_k, y_k) \frac{V_k(x, y; C_n^*)}{V(C_n^*)} + q_n(x, y) - u(x, y) \right| \\ &\leqq \|u - q_n\| \left(1 + \sum_{k=0}^n |V_k(x, y; C_n^*)/V(C_n^*)| \right). \end{aligned}$$

Since

$$\max_{C_n \subset \mathrm{cl}\,(\Omega)} |V(C_n)| = \max_{C_n \subset \partial\Omega} |V(C_n)|,$$

$|V_k(x, y; C_n^*)/V(C_n^*)| \leqq 1$ for every $(x, y) \in$ cl (Ω). Therefore

$$\|u - L_n\| \leqq (n+2)\|u - q_n\| \leqq (n+2)M/R^n.$$

4. Remarks concerning the Dirichlet problem. Over the past 15 years the constructive solution of the Dirichlet problem for Laplace's equation in two dimensions by harmonic polynomial interpolation to the boundary values has been extensively developed (see [1]–[6], [13]). This success suggests the possibility of solving the Dirichlet problem for more general elliptic equations by interpolating to the boundary values with polynomial solutions of the equation. The above results regarding the uniform convergence of GASP polynomial interpolants to a GASP which is regular on the closure of a region should be regarded as a first step in this direction. Before closing, we provide here some initial remarks concerning such a constructive solution of the Dirichlet problem for GASP's.

The solution of the GASP Dirichlet problem for a bounded, axiconvex region Ω having a smooth boundary always exists and is unique [16]. That is, if $v \in C(\partial\Omega)$, $v(x, y) = v(x, -y)$, then for any $\mu > 0$ there exists a unique GASP u which is regular in Ω, continuous on cl (Ω), and equals v on $\partial\Omega$. Thus letting $CS(\partial\Omega) = \{v : v \in C(\partial\Omega)$, and $v(x, -y) = v(x, y)\}$ we have the following result.

THEOREM 4.1. *Let Ω be a bounded axiconvex star shaped region having a smooth boundary. Then the GASP polynomials are dense (uniform convergence) in $CS(\partial\Omega)$.*

Further if $v \in CS(\partial\Omega)$ *and* $\{q_n\}$ *is a sequence of GASP polynomials uniformly convergent to* v *on* $\partial\Omega$, *then* $\{q_n\}$ *converges uniformly on* cl (Ω) *and* lim q_n *provides the solution of the corresponding Dirichlet.*

Proof. The density is an immediate consequence of Theorem 2.4 and the existence of the solution of the Dirichlet problem. The latter claim is an immediate consequence of the existence of the solution of the Dirichlet problem, and the maximum principle.

Using the result of this theorem, solutions of the Dirichlet problem on an ellipse by GASP polynomial interpolation to the boundary values can be obtained as immediate consequences of well known results regarding the convergence of polynomials interpolating a continuous function on an interval. Let E be an ellipse having an axis (either major or minor) along the x-axis. We suppose (for convenience) that this axis is the interval $[-1, 1]$. Let $v(x, y)$ be a continuous function defined on E which is even in y, and $C_n = \{(x_k^{(n)}, y_k^{(n)})\}_{k=0}^n \subset E$ where $x_k^{(n)} \neq x_j^{(n)}$ for $k \neq j$. Then E has equation $x^2 + y^2/b^2 = 1$, and for $(x, y) \in E$, $v(x, y) = v(x, b\sqrt{1-x^2})$. The function $v(x, b\sqrt{1-x^2})$ is continuous on $[-1, 1]$, and since $L_n(x, y; C_n; v)$ is a polynomial of degree n in x, y, and even in y, $L_n(x, b\sqrt{1-x^2}; C_n; v)$ is a polynomial of degree n in x, and interpolates $v(x, b\sqrt{1-x^2})$ at the points $x_j^{(n)} \in [-1, 1]$, $j = 0, \cdots, n$. Further, by Theorem 4.1, the uniform convergence

$$L_n(x, b\sqrt{1-x^2}; C_n; v) \to v(x, b\sqrt{1-x^2}), \qquad x \in [-1, 1], \tag{4.1}$$

is sufficient to insure that the GASP interpolating polynomials $L_n(x, y; C_n; v)$ converge uniformly on the closure of the region interior to E, with limit the corresponding solution of the Dirichlet problem.

The convergence (4.1) has been perhaps the most extensively studied problem in the theory of interpolation and approximation, yielding here a great number of theorems regarding the constructive solution of GASP Dirichlet problem on an ellipse. For example, appealing to the result of Bernstein [15, p. 50] we have the following.

THEOREM 4.2. *Let* $\Omega = \{(x, y): x^2 + y^2/b^2 < 1\}$, *and* $v(x, y)$ *be even in* y *and satisfy a Lipschitz condition of any positive order on* $\partial\Omega$. *If* $C_n = \{(x_k^{(n)}, y_k^{(n)})\}_{k=1}^{n+1} \subset \partial\Omega$, *where*

$$x_k^{(n)} = \cos \frac{2k-1}{(2n+2)}\pi,$$

then $\lim_{n\to\infty} L_n(x, y; C_n; v)$ *provides the solution of the corresponding Dirichlet problem and the convergence is uniform on* cl (Ω).

Proof. It suffices to show $L_n \to v$ uniformly on $\partial\Omega$. Since $v(x, y)$ satisfies a Lipschitz condition of order $\alpha > 0$ on $\partial\Omega$, $f(x) = v(x, b\sqrt{1-x^2})$ satisfies a Lipschitz condition of order $\alpha/2$ on $[-1, 1]$. Thus by Bernstein's theorem, the polynomials of degree n which interpolate f at

$$x_k^{(n)} = \cos \frac{2k-1}{(2n+2)}\pi, \qquad k = 1, 2, \cdots, n+1,$$

converge uniformly to f on $[-1, 1]$. But these are precisely $L_n(x, b\sqrt{1-x^2}; C_n; v)$.

The reader may find it of interest to compare this result (and others which can be obtained similarly) with that of Walsh [21] regarding the solution of the Dirichlet problem on an ellipse for Laplace's equation in two dimensions by harmonic polynomial interpolation to the boundary values.

REFERENCES

[1] J. H. CURTISS, *Interpolation with harmonic and complex polynomials to boundary values*, J. Math. Mech., 9 (1960), pp. 167–192.

[2] ———, *Interpolation by harmonic polynomials*, J. Soc. Indust. Appl. Math., 10 (1962), pp. 709–736.

[3] ———, *Harmonic interpolation in Fejer points with Faber polynomials as a basis*, Math. Z., 86 (1964), pp. 75–92.

[4] ———, *Solutions of the Dirichlet problem in the plane by approximation with Faber polynomials*, SIAM J. Numer. Anal., 3 (1966), pp. 204–228.

[5] ———, *The transfinite diameter and extremal points for harmonic polynomial interpolation*, J. Analyse Math., 17 (1966), pp. 369–382.

[6] ———, *Transfinite diameter and harmonic polynomial interpolation*, Ibid., 22 (1969), pp. 371–389.

[7] ———, *Overdetermined harmonic polynomial interpolation*, J. Approximation Theory, 5 (1972), pp. 149–175.

[8] A. ERDÉLYI, *Singularities of generalized axially symmetric potentials*, Comm. Pure Appl. Math., 9 (1956), pp. 403–414.

[9] A. FRYANT, *Interpolation and approximation of axisymmetric harmonic functions in R^3*, Amer. J. Math., to appear.

[10] R. P. GILBERT, *Function Theoretic Methods in Partial Differential Equations*, Academic Press, New York, 1969.

[11] G. G. LORENTZ, *Approximation of Functions*, Holt, Rinehart and Winston, New York, 1966.

[12] M. MARDEN, *Axisymmetric harmonic interpolating polynomials in R^n*, Trans. Amer. Math. Soc., 196 (1974), pp. 385–402.

[13] K. MENKE, *Lösung des Dirichlet-Problems bei Jordangebieten mit analytischen Rand durch Interpolation*, Monatsh. Math., 80 (1976), pp. 297–306.

[14] B. MUCHENHOUPT AND E. M. STEIN, *Classical expansions and their relation to conjugate harmonic functions*, Trans. Amer. Math. Soc., 118 (1965), pp. 17–92.

[15] I. NATANSON, *Constructive Function Theory*, vol. 3, Fredric Ungar, New York, 1965.

[16] S. V. PARTER, *On the existence and uniqueness of axially symmetric potentials*, Arch. Rational Mech. Anal., 20 (1965), pp. 279–286.

[17] V. I. SMIRNOV AND N. A. LEBEDEV, *Functions of a Complex Variable*, MIT Press, Cambridge, MA, 1968.

[18] J. L. WALSH, *The approximation of harmonic functions by harmonic polynomials and by harmonic rational functions*, Bull. Amer. Math. Soc., 35 (1929), pp. 499–544.

[19] ———, *On interpolation to harmonic functions by harmonic polynomials*, Proc. Nat. Acad. Sci. U.S.A., 18 (1932), pp. 514–517.

[20] ———, *Interpolation and Approximation by Rational Functions in the Complex Domain*, American Mathematical Society, Providence, RI, 1935.

[21] ———, *Solution of the Dirichlet problem for the ellipse by interpolating harmonic polynomials*, J. Math. Mech., 9 (1960), pp. 193–196.

[22] A. WEINSTEIN, *Generalized axially symmetric potential theory*, Bull. Amer. Math. Soc., 59 (1953), pp. 20–38.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0015 $01.00/0

ASYMPTOTIC BOUNDS OF SOLUTIONS OF THE FUNCTIONAL DIFFERENTIAL EQUATION $x'(t)=ax(\lambda t)+bx(t)+f(t)$, $0<\lambda<1$*

ENG-BIN LIM†

Abstract. In this paper, the author studies the asymptotic bounds of the solutions of the function differential equation

$$x'(t)=ax(\lambda t)+bx(t)+f(t)$$

where $0<\lambda<1$, $a\neq 0$ is a complex constant and $b\neq 0$ is a real constant. His results generalize that of Kato and McLeod [4].

1. Introduction. The objective of this paper is to study the asymptotic bounds of the solutions of the nonhomogeneous functional differential equation

$$x'(t)=ax(\lambda t)+bx(t)+f(t) \tag{1.1}$$

where $0<\lambda<1$, $a\neq 0$ is a complex constant, $b\neq 0$ is a real constant and $f(t)$ is a continuous function defined on the interval $[0,\infty)$. The corresponding homogeneous equation

$$x'(t)=ax(\lambda t)+bx(t) \tag{1.2}$$

arises in [6] where Ockendon and Taylor study the motion of a pantograph head on an electric locomotive. Equation (1.2) has been discussed by Kato and McLeod [4].

Throughout this paper we use the notations: $c=\ln\lambda<0$ and $\kappa=(\ln|b/a|)/c$. We summarize our results in the following theorems.

THEOREM 1. *Let $b<0$. Assume that f' exists. Let $f(t)=O(t^\alpha)$ and $f'(t)=O(t^{\alpha-1})$ where α is a real constant. Then*:

(i) *If $\alpha<\kappa$, every solution of* (1.1) *is $O(t^\kappa)$ as $t\to\infty$.*
(ii) *If $\alpha=\kappa$, every solution of* (1.1) *is $O(t^\kappa\ln t)$ as $t\to\infty$.*
(iii) *If $\alpha>\kappa$, every solution of* (1.1) *is $O(t^\alpha)$ as $t\to\infty$.*

THEOREM 2. *Let $b>0$. Let $f(t)=O(t^\alpha)$. Then every solution of* (1.1) *is $O(e^{bt})$ as $t\to\infty$.*

2. Preliminaries. By a solution of (1.1) we mean a complex-valued continuous function $x(t)$ defined in some subinterval of $[0,\infty)$ and satisfying (1.1). Let $t_0>0$ and suppose that the solution $x(t)$ of (1.1) is known for $t\in[\lambda t_0,t_0]$; then $x(\lambda t)$ is known for $t\in[t_0,\lambda^{-1}t_0]$ and the solution can be extended uniquely to the interval $[t_0,\lambda^{-1}t_0]$ by solving (1.1) and requiring that the solution be continuous at $t=t_0$. By continuing this process we see that the solution exists for all $t\geqq t_0$. Thus, given a continuous function $\varphi(t)$ defined in the interval $[\lambda t, t_0]$, there is a unique solution $x(t)$ of (1.1) for $t\geqq t_0$ satisfying the condition $x(t)=\varphi(t)$ for $t\in[\lambda t_0,t_0]$.

DEFINITION 1. Let $x(t)$ and $g(t)$ be complex-valued functions defined on $[0,\infty)$. We say that $x(t)$ is $O\{g(t)\}$ as $t\to\infty$ if there are constants $K>0$ and $N>0$ such that $|x(t)|\leqq K|g(t)|$ for $t\geqq N$.

LEMMA 1. *Let $c<0$, $\gamma>0$, and l be a complex number such that $|l|=e^{\gamma c}$. Let $W(s)$ be a solution of the difference equation*

$$W(s)-lW(s+c)=h(s) \tag{2.1}$$

* Received by the editors September 9, 1976, and in revised form February 22, 1977.
† School of Mathematical Sciences, University of Science of Malaysia, Minden Penang, Malaysia.

where $h(s)$ is a continuous function. Suppose that

$$|h(s)| \leqq K_1 e^{-\beta s} \quad \textit{for } s \geqq s_0 \tag{2.2}$$

where s_0, K_1 and β are positive constants and $|W(s)| \leqq K_2$ for $s \in [s_0 + c, s_0]$ where K_2 is a positive constant. Then there is a constant M depending on γ, β, c, s_0, K_1, and K_2 such that

(i) $|W(s)| \leqq M e^{-\gamma s}$ *for* $s \geqq s_0$, *if* $\gamma < \beta$.
(ii) $|W(s)| \leqq M s\, e^{-\gamma s}$ *for* $s \geqq s_0$, *if* $\gamma = \beta$.
(iii) $|W(s)| \leqq M e^{-\beta s}$ *for* $s \geqq s_0$, *if* $\gamma > \beta$.

Proof. Let $s \geqq s_0$. Then there is a nonnegative integer n and a number $s_1 \in [s_0, s_0 - c]$ such that $s = s_1 - nc$. It is easy to show by induction that (2.1) implies

$$W(s_1 - nc) = l^{n+1} W(s_1 + c) + \sum_{i=0}^{n} h(s_1 - ic) l^{n-i}, \tag{2.3}$$

$$|W(s)| \leqq K_2 e^{\gamma c(n+1)} + K_1 e^{-\beta s_1 + \gamma nc} \sum_{i=0}^{n} e^{(\beta-\gamma)ic}. \tag{2.4}$$

Case (i): $\gamma < \beta$. Since $c < 0$, the series

$$\sum_{i=0}^{\infty} e^{(\beta-\gamma)ic}$$

is convergent. Let $\sum_{i=0}^{\infty} e^{(\beta-\gamma)ic} = K_3$. Then by (2.4),

$$\begin{aligned} |W(s)| &\leqq K_2 e^{\gamma c + \gamma s_1 - \gamma s} + K_1 K_3 e^{-\beta s_1 + \gamma s_1 - \gamma s} \\ &\leqq K_2 e^{\gamma s_0 - \gamma s} + K_1 K_3 e^{-\gamma s}. \end{aligned}$$

Let $M_1 = K_2 e^{\gamma s_0} + K_1 K_3$. Then $|W(s)| \leqq 2M_1 e^{-\gamma s}$ for $s \geqq s_0$.

Case (ii): $\gamma = \beta$. From (2.4),

$$\begin{aligned} |W(s)| &\leqq K_2 e^{\gamma s_0 - \gamma s} + K_1 e^{-\gamma s_1 + \gamma nc}(n+1) \\ &= K_2 e^{\gamma s_0 - \gamma s} + K_1 e^{-\gamma s_1 + \gamma s_1 - \gamma s}\left(\frac{s_1 + c - s}{c}\right) \\ &\leqq K_2 e^{\gamma s_0 - \gamma s} + \frac{K_1}{|c|} s_0 \cdot e^{-\gamma s} + \frac{K_1}{|c|} \cdot s\, e^{-\gamma s} \\ &\leqq K_2 e^{\gamma s_0 - \gamma s} + \frac{2K_1}{|c|} \cdot s\, e^{-\gamma s}. \end{aligned}$$

Let $M_2 = K_2 e^{\gamma s_0}/s_0 + 2K_1/|c|$. Then $|W(s)| \leqq 2M_2 s\, e^{-\gamma s}$.

Case (iii): $\gamma > \beta$. From (2.4), we obtain

$$|W(s)| \leqq K_2 e^{\gamma s_0 - \gamma s} + K_1 e^{-\beta s_1 + \gamma nc} \sum_{i=0}^{n} e^{(\beta-\gamma)ic}. \tag{2.5}$$

Let $K_4 = 1/(e^{(\beta-\gamma)c} - 1)$. Then

$$\begin{aligned} |W(s)| &\leqq K_2 e^{\gamma s_0 - \gamma s} + K_1 K_4 e^{-\beta s_1 + \gamma nc}[e^{(\beta-\gamma)(n+1)c} + 1] \\ &\leqq K_2 e^{\gamma s_0 - \gamma s} + K_1 K_4 e^{(\beta-\gamma)c - \beta s} + K_1 K_4 e^{(\gamma-\beta)s_1 - \gamma s}. \end{aligned}$$

Since $s_1 \leqq s_0 - c$ and $\gamma > \beta$, therefore

$$|W(s)| \leqq K_2 e^{\gamma s_0 - \beta s} + K_1 K_4 e^{(\beta-\gamma)c - \beta s} + K_1 K_4 e^{(\gamma-\beta)(s_0 - c) - \beta s}.$$

Let $M_3 = K_2 e^{\gamma s_0} + K_1 K_4 e^{(\gamma-\beta)(s_0-c)}$. Then

$$|W(s)| \geqq 3M_3 e^{-\beta s} \quad \text{for } s \geqq s_0.$$

The proof of the lemma is completed by taking $M = 2M_1 + 2M_2 + 3M_3$.

3. Proof of Theorem 1. Let $\delta > \max\{\alpha, \kappa\}$. Let $t = e^s$, $W(s) = t^{-\delta}x(t)$. Then $x(t) = e^{\delta s}W(s)$, and

$$\frac{dx}{dt}(t) = \left[\delta e^{\delta s}W(s) + e^{\delta s}\frac{dW}{ds}\right]e^{-s}.$$

Substituting this into (1.1), we obtain

$$\text{(3.1)} \qquad \frac{dW}{ds}(s) + (\delta - b e^s)W(s) = a\lambda^\delta e^s W(s+c) + e^s e^{-\delta s} f(e^s).$$

Since $f(t) = O(t^\alpha)$ there are positive constants K_5 and t_0 such that $|f(t)| \leqq K_5 t^\alpha$ for all $t \geqq t_0$. Let $g(s) = f(e^s) e^{-\delta s}$. Then

$$\text{(3.2)} \qquad \begin{aligned} |g(s)| &\leqq K_5 e^{\alpha s} e^{-\delta s} \\ &= K_5 e^{(\alpha-\delta)s} \quad \text{if } s \geqq \ln t_0. \end{aligned}$$

Equation (3.1) can be written as

$$\text{(3.3)} \quad \frac{d}{ds}[\exp(\delta s - b e^s)W(s)] = a\lambda^\delta e^s \exp(\delta s - b e^s)W(s+c) + \exp(\delta s - b e^s) e^s g(s).$$

Choose $s_0 \geqq \ln t_0$ so large that $\delta - \frac{1}{2}b e^{s_0} > 0$. Then $\delta s - b e^s$ is increasing for any s in $I_m = [s_0 - mc, s_0 - (m+1)c]$, $m = 0, 1, 2, \cdots$. Let

$$M_m = \max\{|W(s)|, s \in I_m\}, \qquad m = 0, 1, 2, \cdots,$$

$$B_m = \max\{M_m, K_5/|b|\}, \qquad m = 0, 1, 2, \cdots.$$

Write $\tau_m = s_0 - (m+1)c$, $m = 0, 1, 2, \cdots$. Let $s \in I_{m+1}$. Integrating (3.3), we obtain

$$\exp(\delta t - b e^t)W(t)\big|_{\tau_m}^{s} = a\lambda^\delta \int_{\tau_m}^{s} e^t \exp(\delta t - b e^t)W(t+c)\,dt + \int_{\tau_m}^{s} e^t \exp(\delta t - b e^t)g(t)\,dt.$$

This can be written as

$$\begin{aligned} W(s) = {} & \exp[(e^s - e^{\tau_m})b - \delta(s - \tau_m)]W(\tau_m) \\ & + a\lambda^\delta \exp(b e^s - \delta s)\int_{\tau_m}^{s} e^t \exp(\delta t - b e^t)W(t+c)\,dt \\ & + \exp(b e^s - \delta s)\int_{\tau_m}^{s} e^t \exp(\delta t - b e^t)g(t)\,dt. \end{aligned}$$

Therefore,

$$\begin{aligned} |W(s)| \leqq {} & M_m \exp[(e^s - e^{\tau_m})b - \delta(s - \tau_m)] \\ & + M_m|a|\lambda^\delta \exp(b e^s - \delta s)\int_{\tau_m}^{s} e^t \exp(\delta t - b e^t)\,dt \\ & + K_5 e^{(\alpha-\delta)\tau_m} \cdot \exp(b e^s - \delta s)\int_{\tau_m}^{s} e^t \exp(\delta t - b e^t)\,dt, \quad \text{by (3.2).} \end{aligned}$$

Since $\delta > \kappa$ and $|a|\lambda^\delta < |b|$, we have

$$
\begin{aligned}
|W(s)| \leqq{} & M_m \exp\left[(e^s - e^{\tau_m})b - \delta(s-\tau_m)\right] \\
& + M_m |b| \exp(b\,e^s - \delta s) \int_{\tau_m}^{s} e^t \exp(\delta t - b\,e^t)\,dt \\
& + \frac{K_5}{|b|} e^{(\alpha-\delta)\tau_m} |b| \exp(b\,e^s - \delta s) \int_{\tau_m}^{s} e^t \exp(\delta t - b\,e^t)\,dt.
\end{aligned}
$$

Thus

$$
\begin{aligned}
|W(s)| \leqq{} & M_m \exp\left[(e^s - e^{\tau_m})b - \delta(s-\tau_m)\right] \\
& + \left(M_m + \frac{K_5}{|b|} e^{(\alpha-\delta)\tau_m}\right) \cdot \exp(b\,e^s - \delta s) \int_{\tau_m}^{s} |b|\, e^t \exp(\delta t - b\,e^t)\,dt.
\end{aligned}
\tag{3.4}
$$

The integral is integrated by parts to give

$$
\left[\frac{-b\,e^t}{\delta - b\,e^t} \exp(\delta t - b\,e^t)\right]\Bigg|_{\tau_m}^{s} + \int_{\tau_m}^{s} \exp(\delta t - b\,e^t) \frac{d}{dt}\left(\frac{b\,e^t}{\delta - b\,e^t}\right) dt.
\tag{3.5}
$$

The expression (3.5) is equal to

$$
\left[\left(1 - \frac{\delta}{b\,e^t}\right)^{-1} \exp(\delta t - b\,e^t)\right]\Bigg|_{\tau_m}^{s} + O\left[\exp(\delta s - b\,e^s) \int_{\tau_m}^{s} \left|\frac{d}{dt}\left(\frac{b\,e^t}{\delta - b\,e^t}\right)\right| dt\right].
$$

Due to the choice of s_0, the O-term in the last expression is uniform in s, δ, s_0 and m. The above expression can be written as

$$
\exp(\delta t - b\,e^t)|_{\tau_m}^{s} + O[\delta \exp(\delta s - b\,e^s) \exp(-\tau_m)].
$$

Substituting this into (3.4) we obtain

$$
\begin{aligned}
|W(s)| \leqq{} & M_m \exp\left[(e^s - e^{\tau_m})b - \delta(s-\tau_m)\right] \\
& + \left(M_m + \frac{K_5}{|b|} e^{(\alpha-\delta)\tau_m}\right) \exp(b\,e^s - \delta s) \Bigg\{ \exp(\delta t - b\,e^t)\Big|_{\tau_m}^{s} \\
& \qquad + O[\delta \exp(\delta s - b\,e^s) \exp(-\tau_m)] \Bigg\}.
\end{aligned}
$$

Thus

$$
\begin{aligned}
|W(s)| \leqq{} & M_m \exp\left[(e^s - e^{\tau_m})b - \delta(s-\tau_m)\right] \\
& + M_m - M_m \exp\left[(e^s - e^{\tau_m})b - \delta(s-\tau_m)\right] \\
& + M_m O(\delta\, e^{-\tau_m}) + \frac{K_5}{|b|} e^{(\alpha-\delta)\tau_m} + \frac{K_5}{|b|} O(\delta\, e^{-\tau_m}) \\
\leqq{} & B_m[1 + O(e^{-\tau_m}) + O(e^{(\alpha-\delta)\tau_m})].
\end{aligned}
$$

Let $r = \min\{1, \delta - \alpha\}$. Then $r > 0$ and

$$
|W(s)| \leqq B_m[1 + O(e^{-r\tau_m})], \qquad s \in I_{m+1}.
$$

This implies that

$$
M_{m+1} \leqq B_m[1 + O(e^{-r\tau_m})] \quad \text{and so} \quad B_{m+1} \leqq B_m[1 + O(e^{-r\tau_m})].
$$

Hence

$$B_m \leqq B_1 \prod_{i=0}^{m-1} [1+O(e^{-r\tau_i})], \qquad m=2,3,\cdots.$$

Since $\tau_i = s_0-(i+1)c$,

$$\prod_{i=0}^{m-1} [1+O(e^{-r\tau_i})] = \prod_{i=0}^{m-1} [1+O(e^{rci})].$$

Thus the convergence of the infinite product

$$\prod_{i=0}^{\infty} [1+O(e^{rci})]$$

implies that B_m is bounded for all m. So M_m is bounded for all m. This proves that $W(s)$ is bounded and $x(t)=O(t^\delta)$.

Differentiating (1.1), we obtain

$$x''(t)=a\lambda x'(\lambda t)+bx'(t)+f'(t).$$

Let δ be the same constant as above. Note that $f'(t)=O(t^{\alpha-1})$. Let

$$\kappa' = (\ln |b/(\lambda a)|)/c.$$

Then $\kappa' = \kappa - 1$. Thus $\delta - 1 > \max\{\alpha-1, \kappa'\}$.

By the above argument, $x'(t)=O(t^{\delta-1})$. Since $W(s)=t^{-\delta}x(t)=e^{-\delta s}x(t)$, therefore $W'(s)=-\delta e^{-\delta s}x(t)+e^{-\delta s}x'(t)\,e^s$.

Since $x(t)=O(t^\delta)$ and $x'(t)=O(t^{\delta-1})$, $W'(s)$ is bounded. Equation (3.1) can be written as

$$W(s)-\left(\frac{a\lambda^\delta}{-b}\right)W(s+c)=\frac{1}{b}e^{-s}[W'(s)+\delta W(s)-e^s g(s)]. \tag{3.6}$$

Let

$$l=\frac{a\lambda^\delta}{-b}, \qquad h(s)=\frac{1}{b}e^{-s}[W'(s)+\delta W(s)-e^s g(s)].$$

Then equation (3.6) can be written as

$$W(s)-lW(s+c)=h(s) \tag{3.7}$$

where

$$|l| = \left|\frac{a\lambda^\delta}{a\lambda^\kappa}\right| = \lambda^{\delta-\kappa}=e^{(\delta-\kappa)c}$$

and there are positive constants A, B and s_0 such that

$$|h(s)| \leqq A\,e^{-s}+B\,e^{(\alpha-\delta)s}, \qquad s \geqq s_0.$$

Case (i): $\alpha<\kappa$. If $\kappa<1+\alpha$, we choose δ such that $\kappa<\delta<1+\alpha$. Then

$$|h(s)| \leqq (A+B)\,e^{(\alpha-\delta)s}, \qquad s\geqq s_0.$$

Let $\gamma=\delta-\kappa$, $\beta=\delta-\alpha$. Then $\gamma<\beta$. Therefore by Lemma 1, there is a constant M such that $|W(s)|\leqq M e^{(\kappa-\delta)s}$ for $s \geqq s_0$. Thus $x(t)=t^\delta W(s)=O(e^{\kappa s})=O(t^\kappa)$ as $t\to\infty$.

If $\kappa \geqq 1+\alpha$, choose $\delta=\kappa+\frac{1}{2}$. Then $|h(s)| \leqq (A+B)\, e^{-s}$, $s \geqq s_0$. Let $\gamma=\delta-\kappa$, $\beta=1$. Then $\gamma<\beta$. By Lemma 1, there is a constant M such that $|W(s)| \leqq M e^{(\kappa-\delta)s}$ for $s \geqq s_0$. Thus $x(t)=O(t^{\kappa})$ as $t \to \infty$.

Case (ii): $\alpha=\kappa$. Choose $\delta=\kappa+\frac{1}{2}$. Then $|h(s)| \leqq (A+B)\, e^{(\alpha-\delta)s}$, $s \geqq s_0$. By Lemma 1, $|W(s)| \leqq Ms\, e^{(\kappa-\delta)s}$ for $s \geqq s_0$. Thus $x(t)=O(t^{\kappa} \ln t)$ as $t \to \infty$.

Case (iii): $\alpha>\kappa$. Choose $\delta=\alpha+\frac{1}{2}$. Then $|h(s)| \leqq (A+B)\, e^{(\alpha-\delta)s}$, $s \geqq s_0$ and $\gamma=\delta-\kappa>\beta=\delta-\alpha$. By Lemma 1, $|W(s)| \leqq M e^{(\alpha-\delta)s}$, $s \geqq s_0$. Thus $x(t)=O(t^{\alpha})$ as $t \to \infty$. This completes the proof of Theorem 1.

4. Proof of Theorem 2. Equation (1.1) can be written in the form

$$\frac{d}{dt}(x(t)\, e^{-bt})=ax(\lambda t)\, e^{-bt}+f(t)\, e^{-bt}. \tag{4.1}$$

Choose t_0 so large that $\lambda bt>\alpha \ln t$ for $t \geqq t_0$. There is a constant $K>0$ such that $|f(t)| \leqq Kt^{\alpha}$, $t \geqq t_0$. Let

$$\tau_m=t_0/\lambda^m, \qquad m=0,1,2,\cdots,$$

$$I_m=[t_0/\lambda^m, t_0/\lambda^{m+1}], \qquad m=0,1,2,\cdots.$$

Let $\tau \in I_{m+1}$ and let $M_m=\sup\{|x(t)\, e^{-bt}| : t \in I_m\}$. Then $|x(t)| \leqq M_m e^{bt}$, $t \in I_m$. Integrating (4.1) and taking modulus, we obtain

$$\begin{aligned}|x(\tau)\, e^{-b\tau}| &\leqq |x(\tau_m)\, e^{-b\tau_m}|+|a|M_m \int_{\tau_m}^{\tau} e^{(\lambda-1)bt}\, dt+K\int_{\tau_m}^{\tau} t^{\alpha} e^{-bt}\, dt \\ &\leqq M_m+M_m \cdot O(e^{(\lambda-1)b\tau_m})+K \cdot O(e^{(\lambda-1)b\tau_m}).\end{aligned}$$

Let $B_m=\max\{K, M_m\}$. Then $|x(\tau)\, e^{-b\tau}| \leqq B_m[1+O(e^{(\lambda-1)b\tau_m})]$. This implies that $M_{m+1} \leqq B_m[1+O(e^{(\lambda-1)b\tau_m})]$, and so $B_{m+1} \leqq B_m[1+O(e^{(\lambda-1)b\tau_m})]$. Hence

$$B_m \leqq B_0 \prod_{i=0}^{m-1} [1+O(e^{(\lambda-1)b\tau_i})].$$

Thus B_m is bounded and so is M_m for all m. Therefore $x(t)=O(e^{bt})$ as $t \to \infty$. This completes the proof of Theorem 2.

REFERENCES

[1] R. E. BELLMAN AND K. L. COOKE, *Differential-difference Equations*, Academic Press, New York, 1963.

[2] L. FOX, D. F. MAYERS, J. R. OCKENDON AND A. B. TAYLOR, *On a functional differential equation*, J. Inst. Math. Appl., 8 (1971), pp. 271–307.

[3] M. L. HEARD, *Asymptotic behavior of solutions of the functional differential equation* $x'(t)=ax(t)+bx(t^{\alpha})$, $\alpha>1$, J. Math. Anal. Appl., 44 (1973), pp. 745–757.

[4] T. KATO AND J. B. MCLEOD, *The functional-differential equation* $y'(x)=ay(\lambda x)+by(x)$, Bull. Amer. Math. Soc., 77 (1971), pp. 891–937.

[5] T. KATO, *Asymptotic behavior of solutions of the functional differential equation* $y'(x)=ay(\lambda x)+by(x)$, Delay and Functional Differential Equations and their Applications, Klaus Schmitt, ed., Academic Press, New York, 1972, pp. 197–217.

[6] J. R. OCKENDON AND A. B. TAYLOR, *The dynamics of a current collection system for an electric locomotive*, Proc. Roy. Soc. London Ser. A, A322 (1971), pp. 447–468.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0016 $01.00/0

MULTILINEAR GENERATING FUNCTION FOR THE KONHAUSER BIORTHOGONAL POLYNOMIAL SETS*

K. R. PATIL†, AND N. K. THAKARE‡

Abstract. The purpose of this note is to obtain a multilinear generating function for the Konhauser biorthogonal polynomials.

1. Introduction. The polynomials

$$Z_n^\alpha(x;k)=\frac{\Gamma(kn+\alpha+1)}{n!}\sum_{j=0}^{n}(-1)^j\binom{n}{j}\frac{x^{kj}}{\Gamma(kj+\alpha+1)},$$

$$Y_n^\alpha(x;k)=\frac{1}{n!}\sum_{r=0}^{\infty}\frac{x^r}{r!}\sum_{s=0}^{r}(-1)^s\binom{r}{s}\left(\frac{s+\alpha+1}{k}\right)_n$$

form a biorthogonal system; see Carlitz [1], Preiser [6] and Konhauser [3], [4]. We shall call $Z_n^\alpha(x;k)$ and $Y_n^\alpha(x;k)$ the biorthogonal polynomials of the first and second kind, respectively.

The operational formula

$$\exp(t\theta)[x^\alpha f(x)]=x^\alpha(1-tkx^k)^{-(\alpha+\lambda)/k}f[x(1-x^kkt)^{-1/k}] \tag{1.1}$$

where $\theta=x^k+x^{k+1}\,d/dx$, follows from formula (1.9) in Mittal [5] by changing the variable x to x^k.

It is easy to see that

$$Y_n^{\alpha+\lambda}(x;k)=\frac{x^{-(\alpha+1+kn)}}{k^n n!}\exp(x)\theta^n[\exp(-x)x^{\alpha+1}]. \tag{1.2}$$

The purpose of this note is to obtain a multilinear generating function for the Konhauser biorthogonal polynomials.

2. Multilinear generating function.

THEOREM. *For the Konhauser biorthogonal polynomial sets* $\{Z_n^\alpha(x;k)\}$ *and* $\{Y_n^\alpha(x;k)\}$ *we have the generating relation*

$$\begin{aligned}\sum_{n_1,n_2,\cdots,n_r=0}^{\infty}&(m+n_1+n_2+\cdots+n_r)!\,Y_{m+n_1+n_2+\cdots+n_r}^{\alpha+\lambda}(x;k)\\&\cdot\prod_{i=1}^{r}[u_i^{n_i}Z_{n_i}^{\beta_i}(y_i;s)/(1+\beta_i)_{sn_i}]\\&=e^x\{\Delta_r\}^{(-(\alpha+1)/k-m)(1+\lambda)}\\&\sum_{l=0}^{\infty}\frac{1}{l!}\left(\frac{\alpha+\lambda+l+1}{k}\right)_m\left(-\frac{x}{\Delta_r^{(\lambda+1)/k}}\right)^l\\&\cdot\psi_2^r\left[\frac{(\alpha+\lambda+l+1)}{k}+m;(1+\beta_1)^s,\cdots,(1+\beta_r)^s;\right.\\&\left.\left(\frac{-y_1^su_1}{s\Delta_r^{\lambda+1}}\right),\cdots,\left(\frac{-y_r^su_1}{s\Delta_r^{\lambda+1}}\right)\right],\end{aligned} \tag{2.1}$$

* Received by the editors June 14, 1976, and in revised form December 31, 1976.

† Arts, Science and Commerce College, Ramanandnagar, District Sangli, India 416 308.

‡ Department of Mathematics, Shivaji University, Kolhapur, India, now at Marathwada University, Aurangabad, India, 431 004.

where ψ_2^n is a confluent hypergeometric function of n variables defined by [2, p. 445]

$$\psi_2^n[a; b_1, b_2, \cdots, b_n; x_1, x_2, \cdots, x_n] = \sum_{m_1,m_2,\cdots,m_n=0}^{\infty} a_{m_1+m_2+\cdots+m_n} \prod_{i=1}^{n} \frac{(x_i)^{m_i}}{(b_i)_{m_i} m_i!}, \tag{2.2}$$

and for convenience we have

$$\Delta_r = 1 - \sum_{i=1}^{r} u_i, \qquad r = 1, 2, 3, \cdots.$$

Proof. Let $I_r(u_1, u_2, \cdots, u_r)$ denote the left-hand side of (2.1). Replace u_i by $u_i kx^k$ $(i = 1, 2, \cdots, r)$ and use the relation (1.2) to obtain

$$I_r(u_1kx^k, u_2kx^k, \cdots, u_rkx^k) = \sum_{n_1,n_2,\cdots,n_r=0}^{\infty} k^{-m} \exp(x) x^{-(\alpha+1+km)} \theta^{m+n_1+\cdots+n_r} \cdot \exp(-x) x^{\alpha+1} \prod_{i=1}^{r} (u_i kx^k)^{n_i} Z_{n_i}^{\beta_i}(y_i; s)/(1+\beta_i)_{sn_i}.$$

One also has (see Rainville [7, formulas (1), (7), (8) on pp. 132, 133])

$$\sum_{n=0}^{\infty} \frac{Z_n^{\alpha}(x; k)t^n}{(1+\alpha)_{sn}} = \exp t\, {}_0F_s\left[-; \frac{\alpha+1}{s}, \frac{\alpha+2}{s}, \cdots, \frac{\alpha+s}{s}; -\left(\frac{x}{s}\right)^s t\right].$$

On account of this we now have

$$\begin{aligned} &k^m \exp(-x) x^{\alpha+1+km} I_r(u_1kx^k, u_2kx^k, \cdots, u_rkx^k) \\ &= \theta^m \prod_{i=1}^{r} \exp(u_i\theta)\, {}_0F_s\left[-; \frac{\beta_i+1}{s}, \frac{\beta_i+2}{s}, \cdots, \frac{\beta_i+s}{s}; -\left(\frac{y_i}{s}\right)^s u_i\theta\right] \exp(-x) x^{\alpha+1}, \\ &= \exp\left(\sum_{i=1}^{r} u_i\theta\right) \sum_{n_1,n_2,\cdots,n_r=0}^{\infty} (-1)^{n_1+n_2+\cdots+n_r} \\ &\quad \cdot \prod_{i=1}^{r} \left[\left(\frac{y_i^s u_i}{s}\right)^{n_i} \Big/ n_i!(1+\beta_i)_{n_i}^s\right] \theta^{m+n_1+n_2+\cdots+n_r} \sum_{l=0}^{\infty} \frac{(-1)}{l!} x^{\alpha+l+1}, \\ &= \sum_{l,n_1,n_2,\cdots,n_r=0}^{\infty} (-1)^{l+n_1+n_2+\cdots+n_r} \exp\left(\sum_{i=1}^{r} u_i\theta\right) x^{1+\alpha+l+k(m+n_1+n_2+\cdots+n_r)} \\ &\quad \cdot \frac{k^{m+n_1+n_2+\cdots+n_r}}{l!} \left(\frac{\alpha+\lambda+l+1}{k}\right)_m \left(\frac{\alpha+\lambda+l+1}{k}+m\right)_{n_1+n_2+\cdots+n_r} \\ &\quad \cdot \prod_{i=1}^{r} \left[\left(\frac{y_1^s u_i}{s}\right)^{n_i} \Big/ n_i!(1+\beta_i)_{n_i}^s\right]. \end{aligned}$$

Using (1.1) we observe that the above expression is equal to

$$\begin{aligned} &\sum_{l,n_1,n_2,\cdots,n_r=0}^{\infty} (-1)^{l+n_1+n_2+\cdots+n_r} \left(\frac{\alpha+\lambda+l+1}{k}\right)_m \left(\frac{\alpha+\lambda+l+1}{k}+m\right)_{n_1+n_2+\cdots+n_r} \\ &\cdot x^{\alpha+l+1+(m+n_1+n_2+\cdots+n_r)k} \frac{k^{m+n_1+n_2+\cdots+n_r}}{l!} \\ &\cdot \prod_{i=1}^{r} \left[\left(\frac{y_i^s u_i}{s}\right)^{n_i} \Big/ n_i!(1+\beta_i)_{n_i}^s\right] \left[1 - \sum_{i=1}^{r} u_i kx^k\right]^{(\lambda+1)\{-(\alpha+l+1)/k-(m+n_1+n_2+\cdots+n_r)\}} \end{aligned}$$

Hence, replacing u_i by u_i/kx^k $(i=1,2,\cdots,r)$ on both sides, we get our final result:

$$I_r(u_1,u_2,\cdots,u_r)=e^x(\Delta_r)^{(\lambda+1)(-(\alpha+1)/k-m)}$$
$$\cdot\sum_{l=0}^{\infty}\left(\frac{\alpha+\lambda+l+1}{k}\right)_m\left(-\frac{x}{\Delta_r^{(\lambda+1)/k}}\right)^l\frac{1}{l!}$$
$$\sum_{n_1,n_2,\cdots,n_r=0}^{\infty}\left(\frac{\alpha+\lambda+l+1}{k}+m\right)_{n_1+n_2+\cdots+n_r}$$
$$\cdot\prod_{i=1}^{r}\left[\left(\frac{-y_i^s u_i}{s\Delta_r^{\lambda+1}}\right)\Big/ n_i!(1+\beta_i)_{n_i}^s\right].$$

This completes the proof.

3. Particular cases.

1. Put $\lambda=0$, $k=s=1$ in (2.1) to get the relation obtained by Srivastava and Singhal [8].

2. Put $\lambda=0$, $k=s=r=1$ in (2.1) to have a bilinear generating function for the generalized Laguerre polynomials, namely

$$\sum_{n=0}^{\infty}\frac{(m+n)!}{(1+\beta)_n}L_{m+n}^{\alpha}(x)L_n^{\beta}(y)z^n$$
$$=(1+\alpha)_m(1-z)^{-\alpha-m-1}e^x\psi_2^2\left[\alpha+m+1;\alpha+1,\beta+1;\frac{-xz}{1-z},\frac{-yz}{1-z}\right].$$

In this context see Erdélyi et al. [2].

Acknowledgment. The authors are grateful to the referees for inviting their attention to the work of Mittal and also for various suggestions which led to the improvement of the paper.

REFERENCES

[1] L. CARLITZ, *A note on certain bi-orthogonal polynomials*, Pacific J. Math., 24 (1968), pp. 425–430.
[2] A. ERDÉLYI, W. MAGNUS, F. OBERHETTINGER AND F. G. TRICOMI, *Tables of Integral Transforms*, vol. II, McGraw-Hill, New York, 1954.
[3] J. D. E. KONHAUSER, *Some properties of bi-orthogonal polynomials*, J. Math. Anal. Appl., 11 (1965), pp. 242–260.
[4] ———, *Biorthogonal polynomials suggested by the Laguerre polynomials*, Pacific J. Math., 21 (1967), pp. 303–314.
[5] H. B. MITTAL, *Operational representation for the generalized Laguerre polynomial*, Glasnik Mat. Ser. III, 6 (1971), pp. 45–53.
[6] S. PREISER, *An investigation of bi-orthogonal polynomials derivable from ordinary differential equations of the third order*, J. Math. Anal. Appl., 4 (1962), pp. 38–64.
[7] EARL D. RAINVILLE, *Special Functions*, Macmillan, New York, 1967.
[8] H. M. SRIVASTAVA AND J. P. SINGHAL, *Some formulas involving the product of several Jacobi or Laguerre polynomials*, Bull. de la classe des Sci. LVIII, 1972, pp. 1238–1247.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0017 $01.00/0

ON A GENERALIZED EULER-POISSON-DARBOUX EQUATION*

D. W. BRESTERS†

Abstract. Solutions of the following Cauchy problem are obtained by means of Fourier transform methods:

$$\sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2} - \frac{\partial^2 u}{\partial t^2} - \frac{\lambda}{t}\frac{\partial u}{\partial t} = c^2 u,$$

$$\left.\begin{array}{l} u(x,t)=f(x) \\ u_t(x,t)=0 \end{array}\right\} \quad \text{at } t=0.$$

Special cases for the parameters λ and c are considered and the regularity of the solutions is studied in some detail.

1. Introduction. In a previous paper ([3]) the author investigated the Cauchy problem

$$\Delta u - \frac{\partial^2 u}{\partial t^2} - \frac{\lambda}{t}\frac{\partial u}{\partial t} = 0,$$

$$(1.1) \qquad \left.\begin{array}{l} u(x,t)=f(x), \\ u_t(x,t)=0 \end{array}\right\} \quad \text{at } t=0,$$

where $x=(x_1, \cdots, x_n)$, $\Delta u = \sum_{i=1}^{n} \partial^2 u/\partial x_i^2$ and λ is a parameter.

A method was developed which enabled us to obtain solutions for all values of the parameter λ, including the so-called exceptional values $\lambda = -1, -3, -5, \cdots$. The results of [3] will be used freely in this paper and for further information the reader is quite often referred to [3]. The present research is concerned with the solution of the following problem:

$$\Delta u - \frac{\partial^2 u}{\partial t^2} - \frac{\lambda}{t}\frac{\partial u}{\partial t} = c^2 u,$$

$$\left.\begin{array}{l} u(x,t)=f(x), \\ u_t(x,t)=0 \end{array}\right\} \quad \text{at } t=0$$

where λ and c are parameters. This problem was studied earlier by Young [11]. Following Weinstein [10] he first obtains solutions for $\lambda \geqq n$. Solutions for other values of λ are then found by means of recurrence relations.

In this paper we use a Fourier method. For $\lambda \geqq n$ this was done first by Carroll [4] (also cf. Carroll and Showalter [5]). He obtained expressions which are identical to those found in our § 2.2. It shall appear, however, that the same method, which uses Fourier transforms with respect to the space variables only, can give us solutions for all values of λ, including the exceptional ones $\lambda = -1, -3, -5, \cdots$. The case in which $\lambda \neq -1, -3, -5, \cdots$ is discussed in § 2. The solution will be found as a generalized function defined on S, the space of testing functions which decrease, together with all their derivatives, faster than any negative power of $|x|$ as $|x| \to \infty$. Here $|x|$ denotes the Euclidean norm $|x| = (x_1^2 + \cdots + x_n^2)^{1/2}$. The variable t is considered as a parameter. We shall need several results on these generalized functions (distributions). Most of

* Received by the editors November 4, 1976, and in revised form September 27, 1977.
† Institute of Applied Mathematics, University of Amsterdam, Amsterdam 1004, The Netherlands.

them may be found in the well-known book by Gel'fand and Shilov [6]. For some rather special results on Fourier transforms the reader is referred to [2].

In § 2.1 we show that for $\lambda > n$ our distributional solution is regular; i.e. it can be expressed by means of convergent integrals. We shall show that in this case our solution is equivalent to the one given by Young. In § 2.2 we obtain an explicit form for the solution for arbitrary $\lambda \neq -1, -3, -5, \cdots$. Next we show that for $c = 0$ one obtains again the results of [3] while for $\lambda = 0$ one obtains the solution of a Cauchy problem for the Klein–Gordon equation. In § 2.4 we discuss the regularity of the solution. Finally, in § 3, we give a brief discussion of the solution in case λ takes one of the exceptional values $-1, -3, \cdots$.

2. Solution of the Cauchy problem.

2.0. General description of the method. The problem under consideration is to find distributional solutions $u(t, x; \lambda)$ of the initial value problem

$$\begin{aligned} &\left(\Delta - \frac{\partial^2}{\partial t^2} - \frac{\lambda}{t}\frac{\partial}{\partial t}\right) u(t, x; \lambda) = c^2 u(t, x; \lambda), \\ &u(0, x; \lambda) = f(x), \\ &u_t(0, x; \lambda) = 0. \end{aligned} \tag{2.1}$$

As in the case $c = 0$ we perform Fourier transformation with respect to x only. We obtain

$$\begin{aligned} &\left(k^2 + c^2 + \frac{\partial^2}{\partial t^2} + \frac{\lambda}{t}\frac{\partial}{\partial t}\right) \hat{u}(t, k; \lambda) = 0, \\ &\hat{u}(0, k; \lambda) = \hat{f}(k), \\ &\hat{u}_t(0, k; \lambda) = 0 \end{aligned} \tag{2.2}$$

where $k = (k_1, \cdots, k_n)$ and $k^2 = k_1^2 + \cdots + k_n^2$.

First we construct a solution $\hat{G}_\lambda(t, k)$ of (2.2) which satisfies the differential equation and the initial data

$$\hat{G}_\lambda(0, k) = 1,$$

$$\frac{\partial}{\partial t}\hat{G}_\lambda(t, k) = 0 \quad \text{at } t = 0.$$

Once we have found $\hat{G}_\lambda(t, k)$ the solution of (2.2) is given by

$$\hat{u}(t, k; \lambda) = \hat{G}_\lambda(t, k) \cdot \hat{f}(k) \tag{2.3}$$

and the solution of (2.1) shall be

$$\begin{aligned} u(t, x; \lambda) &= F_k^{-1}[\hat{G}_\lambda(t, k)] * f(x) \\ &= G_\lambda(t, x) * f(x) \end{aligned} \tag{2.4}$$

where F_k^{-1} denotes the inverse Fourier transform with respect to $k = (k_1, \cdots, k_n)$, and the symbol $*$ denotes convolution with respect to $x = (x_1, \cdots, x_n)$ only.

We call $G_\lambda(t, x)$ the "fundamental solution" of problem (2.1). In order to prove that (2.4) actually represents the solution of (2.1) we shall have to show later on that

(i) $G_\lambda(t, x)$ is a well defined distribution in S'_x, the dual of the space S_x containing testing functions depending on x;

(ii) $G_\lambda(t, x)$ is twice continuously differentiable with respect to the parameter t for all $t \geqq 0$;

(iii) the convolution product in (2.4) exists for a large enough class of functions and distributions $f(x)$.

By means of the transformation

$$\hat{G}_\lambda(t, k) = t^{(1-\lambda)/2} h(t, k)$$

the differential equation in (2.2) becomes a Bessel differential equation which for $\lambda \neq -1, -3, -5, \cdots$ has the following solutions satisfying the initial conditions $\hat{G}_\lambda(t, k) = 1$ and $(\partial/\partial t)\hat{G}_\lambda(t, k) = 0$ for $t = 0$:

$$\begin{aligned} \hat{G}_\lambda(t, k) = {} & 2^{(\lambda-1)/2}\Gamma\left(\frac{\lambda+1}{2}\right) \cdot \{(k^2+c^2)^{1/2}t\}^{(1-\lambda)/2} J_{(\lambda-1)/2}\{(k^2+c^2)^{1/2}t\} \\ & + A t^{(1-\lambda)/2} J_{(1-\lambda)/2}\{(k^2+c^2)^{1/2}t\}. \end{aligned} \tag{2.5}$$

Here A denotes an arbitrary complex number which vanishes for $\lambda \geqq 0$. In the case in which $\lambda = -1, -3, -5, \cdots$ we obtain

$$\begin{aligned} \hat{G}_\lambda(t, k) = {} & \frac{-\pi 2^{(\lambda-1)/2}}{\Gamma((1-\lambda)/2)} \{(k^2+c^2)^{1/2}t\}^{(1-\lambda)/2} Y_{(1-\lambda)/2}\{(k^2+c^2)^{1/2}t\} \\ & + B t^{(1-\lambda)/2} J_{(1-\lambda)/2}\{(k^2+c^2)^{1/2}t\} \end{aligned} \tag{2.6}$$

where B denotes an arbitrary complex number.

From (2.5) and (2.6) it is evident that we have a unique solution for $\lambda \geqq 0$ only. It is also clear that for the "exceptional values" of the parameter λ we shall have a solution of a different character. This is due to the appearance of the Bessel function of the second kind $Y_\nu(z)$ in (2.6).

The problem is thus to find the inverse transform $F_k^{-1}[\hat{G}_\lambda(t, k)]$. In this section we restrict ourselves to the case in which $\lambda \neq -1, -3, -5, \cdots$. The inverse transform mentioned above is then most easily found by considering c as an additional independent variable.

From the results of [3] we obtain, for $\lambda \geqq 0$,

$$F_{k,c}^{-1}[\hat{G}_\lambda(t, k)] = t^{1-\lambda} \frac{\Gamma((\lambda+1)/2)}{\pi^{(n+1)/2}} \frac{(t^2-x^2-\sigma^2)_+^{(\lambda-n-1)/2}}{\Gamma((\lambda-n)/2)} \tag{2.7}$$

where $x^2 = x_1^2 + \cdots + x_n^2$ and σ is the variable corresponding to c. As a consequence we have

$$\begin{aligned} G_\lambda(t, x) &= F_\sigma[F_{k,c}^{-1}[\hat{G}_\lambda(t, k)]] \\ &= \frac{t^{1-\lambda}\Gamma((\lambda+1)/2)}{\pi^{(n+1)/2}} F_\sigma\left[\frac{(t^2-x^2-\sigma^2)_+^{(\lambda-n-2)/2}}{\Gamma((\lambda-n)/2)}\right]. \end{aligned} \tag{2.8}$$

It is easily seen that $G_\lambda(t, x)$ as defined by (2.8) is a well-defined distribution on S_x, the space of testing functions depending on x. The variable t is considered as a parameter. Consequently we have that our original problem (2.1) will, for $\lambda \geqq 0$, be solved by

$$\begin{aligned} u(t, x; \lambda) &= G_\lambda(t, x) * f(x) \\ &= \frac{t^{1-\lambda}\Gamma((\lambda+1)/2)}{\pi^{(n+1)/2}} F_\sigma\left[\frac{(t^2-x^2-\sigma^2)_+^{(\lambda-n-2)/2}}{\Gamma((\lambda-n)/2)}\right] * f(x) \end{aligned} \tag{2.9}$$

where $*$ denotes convolution with respect to x only.

In § 2.4 we shall see that for a large class of initial data $f(x)$, (2.9) is a well-defined distribution which represents the unique solution to the Cauchy problem (2.1) for all $\lambda \geqq 0$. For a more explicit form of the solution we obviously need the Fourier transform

$$F_\sigma\left[\frac{(t^2-x^2-\sigma^2)_+^{(\lambda-n-2)/2}}{\Gamma((\lambda-n)/2)}\right]. \tag{2.10}$$

However, before embarking on the calculation of (2.10) we show first in the next subsection that for $\lambda > n$ the solution can be expressed as a classical integral which is equivalent to the expression found by Young [11].

2.1. The solution for $\lambda > n$. In order to make our results comparable to those of other authors we introduce the symbol ω_s for the surface of the unit sphere in R_s. The introduction of this symbol is due to the fact that most authors follow the Weinstein method which contains as a first step the solution in the case where $\lambda > n$, λ integer. The solution is then found as a mean value of the initial data over a multidimensional sphere. In our method the introduction of ω_s is fully arbitrary and amounts to putting

$$\omega_s = \frac{2\pi^{s/2}}{\Gamma(s/r)} \qquad (s>0).$$

From (2.9) we then find

$$u(t,x;\lambda;c) = \frac{\omega_{\lambda-n}}{\omega_{\lambda+1}} t^{1-\lambda} F_\sigma[(t^2-x^2-\sigma^2)_+^{(\lambda-n-2)/2}] * f(x),$$

which equals, for $\lambda > n$,

$$\frac{\omega_{\lambda-n}}{\omega_{\lambda+1}} \int_{-\infty}^{+\infty} t^{1-\lambda} F_\sigma[(t^2-x^2-\sigma^2)_+^{(\lambda-n-2)/2}] f(x+\xi)\,d\xi$$

where we denoted $(\xi_1, \xi_2, \cdots, \xi_n)$ by ξ. Here we assumed that

$$F_\sigma[(t^2-x^2-\sigma^2)_+^{(\lambda-n-2)/2}]$$

is a regular distribution. It is easily seen that this assumption actually holds if $\lambda > n$.

Putting $\xi = \eta t$ and $\sigma = \nu t$ (and hence $d\xi = t^n\,d\eta$), we obtain

$$\begin{aligned} u(t,x;\lambda;c) &= \frac{\omega_{\lambda-n}}{\omega_{\lambda+1}} \int_{-\infty}^{+\infty} F_\nu[(1-\eta^2-\nu^2)_+^{(\lambda-n-2)/2}] f(x+\eta t)\,d\eta \\ &= 2\frac{\omega_{\lambda-n}}{\omega_{\lambda+1}} \int_{-\infty}^{+\infty} f(x+\eta t) \int_0^\infty (1-\eta^2-\nu^2)_+^{(\lambda-n-2)/2} \cos c\nu t\,d\nu\,d\eta, \end{aligned}$$

or equivalently

$$\begin{aligned} u(t,x;\lambda;c) &= \frac{\omega_{\lambda-n}}{\omega_{\lambda+1}} \iint_{\eta^2+\nu^2<1} f(x+\eta t)(1-\eta^2-\nu^2)^{(\lambda-n-2)/2} \cos c\nu t\,d\nu\,d\eta \\ &= 2\frac{\omega_{\lambda-n}}{\omega_{\lambda+1}} \int_{\eta^2<1} f(x+\eta t) \int_0^{(1-\eta^2)^{1/2}} (1-\eta^2-\nu^2)^{(\lambda-n-2)/2} \cos c\nu t\,d\nu\,d\eta. \end{aligned} \tag{2.11}$$

If, in the last result, we replace c by ic, this is exactly the solution as given for $\lambda > n$ by Young. The integral with respect to ν, appearing in (2.11), obviously converges for $\lambda > n$ only.

It is possible to obtain solutions for other values of λ by means of analytic continuation with respect to λ. This is the general method for extending the domain of definition of certain integrals depending on a parameter. It is described e.g. by Gel'fand and Shilov in [6; Chap. I, § 3]. As an example we take the case where $\lambda = n$. The integral (2.11) in that case is no longer convergent. However, using the fact that the distribution

$$(1-\eta^2-\nu^2)_+^{(\lambda-n-2)/2}$$

has a first order pole at $(\lambda-n-2)/2=-1$ (i.e. $\lambda = n$) with residue

$$\delta(1-\eta^2-\nu^2)$$

we easily obtain that

$$u(t,x;n;c)=\frac{2}{\omega_{n+1}}\int_{\eta^2\leqq 1} f(x+\eta t)\frac{\cos(c(1-\eta^2)^{1/2}t)}{(1-\eta^2)^{1/2}}\,d\eta. \tag{2.12}$$

Using quite different methods Young arrives at the same result (cf. [11]). In principle one may in this way obtain solutions for all values of λ. However, instead of using analytical continuation to extend the integral representation (2.11), we prefer to follow the more direct way of calculating the required Fourier transforms by means of the tables available e.g. in [6]. It should be remarked, however, that this is essentially the same method since the abovementioned tables were constructed by means of analytic continuation.

2.2. General solution for $\lambda \neq -1, -3, -5, \cdots$. From the table of Fourier transforms as given in [4] we use the following:

$$F[(1-x^2)_+^\nu; s] = \pi^{1/2}\Gamma(\nu+1)\left(\frac{s}{2}\right)^{-\nu-1/2} J_{\nu+1/2}(s)$$

where $\nu \neq -1, -2, -3, \cdots$. If we use the formula

$$\left.\frac{(1-x^2)_+^\nu}{\Gamma(\nu+1)}\right|_{\nu=-l} = \delta^{(l-1)}(1-x^2), \qquad l=1,2,\cdots,$$

we may generalize this formula to one valid for all values of ν:

$$F\left[\frac{(1-x^2)_+^\nu}{\Gamma(\nu+1)}; s\right] = \pi^{1/2}\left(\frac{s}{2}\right)^{-\nu-1/2} J_{\nu+1/2}(s).$$

Using a suitable transformation we arrive after some elementary calculations at

$$F_\sigma\left[\frac{(t^2-x^2-\sigma^2)_+^{(\lambda-n-2)/2}}{\Gamma((\lambda-n)/2}\right] = \pi^{1/2}\left(\frac{c}{2\rho}\right)^{(n-\lambda+1)/2} J_{(\lambda-n-1)/2}(\rho c), \tag{2.13}$$

where $\rho=(t^2-x^2)_+^{1/2}$ and $J_\nu(z)$ is the Bessel function of order ν. Consequently we obtain from (2.9) and (2.13) the following formula for the solution $u(t,x;\lambda;c)$ of the initial value problem (2.1), valid for $\lambda \geqq 0$:

(2.14)
$$u(t,x;\lambda;c)=\left(\frac{2}{c}\right)^{(\lambda-n-1)/2}\cdot\frac{\Gamma((\lambda+1)/2)}{\pi^{n/2}}\cdot t^{1-\lambda}\cdot\rho^{(\lambda-n-1)/2}\cdot J_{(\lambda-n-1)/2}(\rho c) * f(x).$$

For $\lambda \geqq n$ Carroll obtained the same result (cf. [4, (4.4)]). A special case ($c=1$) for more general values of λ has been considered by Ossicini [8], [9].

It is easily seen (also cf. [3]) that (2.14) gives the *unique* solution for $\lambda \geqq 0$. In the case in which $\lambda<0$ but $\lambda \neq -1, -3, -5, \cdots$, the situation can be described as follows.

We have as a consequence of (2.5) that (2.14) still gives us a solution. However, we may add any expression of the form

$$t^{1-\lambda}u^*(t, x; 2-\lambda; c)$$

and still have a solution. Here

$$u^*(t, x; 2-\lambda; c)$$

is an arbitrary solution of

$$\left(\Delta-\frac{\partial^2}{\partial t^2}-\frac{2-\lambda}{t}\frac{\partial}{\partial t}\right)u^*=c^2u^*$$

satisfying one initial condition:

$$u^*(0, x; 2-\lambda; c)=0.$$

For details the reader is referred to [3].

2.3. Relation to other Cauchy problems. First of all we check whether for $c\to 0$ we indeed obtain the solution of the Cauchy problem for the ordinary E.P.D. equation which we found in [3]. We use the formula

$$\lim_{z\to 0} z^{-\nu}J_\nu(z)=\frac{1}{2^\nu\Gamma(\nu+1)}, \qquad \nu\neq -1, -2, \cdots,$$

to obtain

$$\lim_{c\to 0} c^{-\nu}J_\nu(c\rho)=\frac{\rho^\nu}{2^\nu\Gamma(\nu+1)}. \tag{2.15}$$

Consequently

$$\lim_{c\to 0} u(t, x; \lambda; c)=t^{1-\lambda}\frac{(t^2-x^2)_+^{(\lambda-n-1)/2}\Gamma((\lambda+1)/2)}{\pi^{n/2}\Gamma((\lambda-n+1)/2)}*f(x)=u(t, x; \lambda; 0) \tag{2.16}$$

for $\lambda\neq -1, -3, -5, \cdots$ and $(\lambda-n-1)/2\neq -1, -2, \cdots$. The condition $(\lambda-n-1)/2\neq -1, -2, \cdots$ can be deleted if we use the formula

$$\lim_{c\to 0}\left(\frac{\rho}{c}\right)^{-l}J_{-l}(\rho c)=\left.\frac{(t^2-x^2)_+^\nu}{2^\nu\Gamma(\nu+1)}\right|_{\nu=-l}$$
$$=2^l\delta^{(l-1)}(t^2-x^2), \qquad l=1, 2, 3, \cdots.$$

In this way we obtain complete agreement with the results of [3].

More interesting is the case in which $\lambda=0$. We should then obtain solutions of the following initial value problem for the Klein–Gordon equation:

$$\Delta v-\frac{\partial^2 v}{\partial t^2}=c^2v$$

$$\left.\begin{aligned} v(t, x)&=f(x),\\ v_t(t, x)&=0\end{aligned}\right\}\quad \text{at } t=0.$$

By letting λ tend to zero in (2.10) we easily obtain

$$v(x, t)=\left(\frac{2}{c}\right)^{(-n-1)/2}\pi^{(1-n)/2}t\rho^{(-1-n)/2}J_{(-n-1)/2}(\rho c)*f(x)$$

with $\rho=(t^2-x^2)_+^{1/2}$. Clearly the first factor in the convolution is concentrated within the light cone as it ought to be.

The behavior near the cone $t^2-x^2=0$ will clearly depend on n being even or odd. More information on this behavior is obtained in the following way.

From the definition of the distributions

$$(Q_+)^{\nu/2}J_\nu(Q_+)$$

where Q is an arbitrary quadratic form (cf. [6, p. 275]) we derive that

$$\rho^\nu J_\nu(\rho)=\sum_{j=0}^{\infty}\frac{(-1)^j(t^2-x^2)_+^{j+\nu}}{j!\,2^{2j+\nu}\Gamma(\nu+j+1)} \tag{2.17}$$

in case $\nu\neq-1,-2,-3,\cdots$. If, however, ν is a negative integer some care is needed. One might be inclined to use the formula

$$J_{-k}(z)=(-1)^kJ_k(z) \qquad (k \text{ pos. integer}) \tag{2.18}$$

and thus arrive at

$$z^{-k}J_{-k}(z)=(-1)^kz^{-k}J_k(z)=\sum_{j=0}^{\infty}\frac{(-1)^{j+k}z^{2j}}{j!\,2^{2j+k}\Gamma(k+j+1)}. \tag{2.19}$$

Substituting $(t^2-x^2)_+^{1/2}$ for z in (2.15), one arrives at the conclusion that the distribution in (2.17) is completely regular for $\nu=-1,-2,-3$. However, this substitution is not allowed because formula (2.14) does not hold in the case where $z=(t^2-x^2)_+^{1/2}$.

If z is a real variable, (2.18) is derived by deleting terms containing $1/(\Gamma(-k+j+1))$ at the poles of $\Gamma(-k+j+1)$, i.e. at $j=1,2,\cdots,k-1$. If, however, z is replaced by $(t^2-x^2)_+^{1/2}$, terms containing

$$\frac{(t^2-x^2)_+^{-k+j}}{\Gamma(-k+j+1)}$$

appear in (2.17). In this case both the numerator and the denominator have a pole at $j=1,2,\cdots$, and consequently these terms cannot be deleted but give rise to δ-type singularities. It is then easily seen that (2.19) should be replaced by

$$\rho^{-k}J_{-k}(\rho)=\sum_{j=0}^{k-1}\frac{(-1)^j}{j!2^{2j-k}}\delta^{(j)}(t^2-x^2)+(-1)^k\rho^{-k}J_k(\rho), \qquad k=1,2,3,\cdots. \tag{2.20}$$

The second part of (2.20) is identical with the right-hand side of (2.19) in case $z=\rho=(t^2-x^2)^{1/2}$. It may be considered as the regular part of the distribution $\rho^{-k}J_{-k}(\rho)$, $k=1,2,\cdots$. As a consequence we now easily obtain for *n even*

$$v(t,x)=\left(\frac{2}{c}\right)^{(-n-1)/2}\pi^{(1-n)/2}t\rho^{(-n-1)/2}J_{(-n-1)/2}(\rho c)*f(x), \tag{2.21}$$

where the singularities on the light cone $t^2-x^2=0$ are of the power type. The highest order singularity in this case is a term containing

$$(t^2-x^2)_+^{(-n-1)/2}*f(x).$$

In case n is odd we obtain

$$v(x,t)=\left(\frac{2}{c}\right)^{(-n-1)/2}\pi^{(1-n)/2}t\cdot\left\{\rho^{(-n-1)/2}J_{(n+1)/2}(\rho c)+\sum_{j=0}^{(n-1)/2}\frac{(-1)^j\delta^{(j)}}{j!\,2^{2j+(n+1)/2}}(t^2-x^2)\right\}*f(x),$$

the highest order singularity thus being found in the term containing

$$\rho^{(n-1)/2}(t^2+x^2)*f(x).$$

It should be remarked that the distribution which in (2.21) and (2.22) is convoluted with $f(x)$ is not the "fundamental solution" of the Cauchy problem as it is usually defined but its derivative. The fundamental solution of the Cauchy problem is defined as that particular elementary solution to our differential operator which vanishes for $t<0$ (cf. [7] or [2]).

After some rather tedious calculations it may then be shown that there is a complete agreement between the results obtained above and those obtained previously by the author for the Klein–Gordon equation (cf. [2, Chap. III]).

2.4. Existence and regularity. It is easily seen that for all values of λ except -1, $-3, -5, \cdots$ the distribution

$$\frac{\Gamma((\lambda+1)/2)}{\pi^{n/2}}t^{1-\lambda}\rho^{(\lambda-n-1)/2}J_{(\lambda-n-1)/2}(\rho c) \tag{2.22}$$

has its support within the sphere $|x|=t$. (As we saw in § 2 we consider t as a parameter taking positive values only.) Consequently the convolution in (2.10) exists in the distributional sense for arbitrary $f(x)\in S'_x$ and the existence of $u(t,x;\lambda;c)$ in the distributional sense is guaranteed for all $t\geqq 0$. We now investigate whether necessary and sufficient conditions on $f(x)$ can be given such that $u(t,x;\lambda;c)$ is a classical solution to our Cauchy problem, i.e. a twice continuously differentiable function with respect to t and x which satisfies (2.1). The differentiability with respect to t is obvious for all $t>0$ and arbitrary initial conditions. The differentiability with respect to x can be investigated as follows.

The relevant part of (2.10) for considering the regularity is the part containing the highest order singularity. This obviously is

$$\frac{(t^2-x^2)_+^{(\lambda-n-1)/2}}{\Gamma((\lambda-n+1)/2)}*f(x) \tag{2.23}$$

and a necessary and sufficient condition for regularity will be that (2.23) is a twice continuously differentiable function of x. In order to obtain conditions on $f(x)$ for (2.23) to be twice continuously differentiable we consider the following distribution:

$$\frac{\partial^m}{\partial x_i^m}(t^2-x^2)_+^\nu*f(x),\qquad \nu\neq-1,-2,\cdots, \tag{2.24}$$

where $m=[\nu+1]=$ the largest integer smaller than or equal to $\nu+1$. In case m is negative the symbol $\partial^m/\partial x_i^m$ denotes the primitive of order $-m$ with respect to x_i. The distribution (2.24) can be written as

$$\frac{\partial^m}{\partial x_i^m}\{\theta(t^2-x^2)(t^2-x^2)^\nu\}*f(x),$$

from which it is easily seen that its most singular part is

$$\delta(t^2-x^2)*f(x). \tag{2.25}$$

The latter is defined and a continuous function of x whenever $f(x)\in C^0$. Consequently (2.24) exists in the classical sense for $f(x)\in C^0$ and will be a C^2-function whenever $f(x)\in C^2$. As a further consequence we obviously have that

$$(t^2-x^2)_+^{\nu}*f(x)\in C^2 \quad \text{if } f(x)\in C^{2-m}.$$

In the case of the distribution (2.23) we obtain that it is regular and belongs to C^2 if we require that $f(x)$ be not less than $2-[\lambda-n+1/2]$ times continuously differentiable. In the case where $(\lambda-n-1)/2$ is a negative integer the same result is even more easily found.

Young [11] obtained that it is sufficient for $f(x)$ to have not less than $(n-\lambda+4)/2$ continuous derivatives. Comparing with our results, we easily see that Young gave indeed sufficient conditions. In the case in which $n-\lambda$ is an odd integer there is a disagreement between the two conditions and the Young condition on $f(x)$ is one too high.

For $\lambda=0$ our results on regularity agree with the ones we obtained on the Klein–Gordon equation in [2]. We there found that $f(x)$ should be not less than $2+[n/2]$ times continuously differentiable.

We now obtain that $f(x)$ should be not less than $2-[(1-n)/2]$ times continuously differentiable. Though these expressions are certainly not equivalent, they yield the same results in our case because n takes positive integer values only.

Again the conditions which Young derives for the solution of the damped wave equation to be a classical solution are sufficient but, in case n is odd, not necessary.

3. The exceptional values of λ. If λ takes one of the values $-1, -3, -5, \cdots$, we obtain solutions which are quite different from those found in the previous section. For $c=0$ we investigated these exceptional cases in detail in [3]. As for all negative values of λ we have that the solution is not uniquely determined. Using the arbitrariness described in § 2.2 we may find the most convenient form for the fundamental solution $G_\lambda(t,x)$ (cf. [3]). We obtain:

$$G_\lambda(t,x)=\frac{\Gamma((n-\lambda+2)/2)}{\Gamma((1-\lambda)/2)}\frac{e^{\pm(n+1)\pi i/2}}{\pi^{(n+1)/2}}t^{1-\lambda}F_\sigma[(t^2-x^2-\sigma^2\pm i0)^{(\lambda-n-2)/2}]. \tag{3.1}$$

For details, the reader is referred to [3, § 4]. Contrary to the situation in the previous section the fundamental solution is no longer concentrated within the sphere $|x|=t$. As in [3] one may show, however, that the convolution

$$u(t,x;\lambda;c)=G_\lambda(t,x)*f(x) \tag{3.2}$$

exists for all $f(x)\in S'$ and $t>0$.

A more explicit form for the fundamental solution $G_\lambda(t,x)$ will depend on the region under consideration. In the interior of the sphere $|x|=t$ we find from the tables of Fourier transforms in [6] that $G_\lambda(t,x)$ is any linear combination of the distributions

$$G_\lambda^{(1)}(t,x)=t^{1-\lambda}\rho^{(\lambda-n-1)/2}H^{(1)}_{(n+1-\lambda)/2}(\rho c)$$

and

$$G_\lambda^{(2)}(t,x)=t^{1-\lambda}\rho^{(\lambda-n-1)/2}H^{(2)}_{(n+1-\lambda)/2}(\rho c)$$

or equivalently

$$(3.3)\quad G_\lambda(t,x)=At^{1-\lambda}\rho^{(\lambda-n-1)/2}J_{(n+1-\lambda)/2}(\rho c)+Bt^{1-\lambda}\rho^{(\lambda-n-1)/2}Y_{(n+1-\lambda)/2}(\rho c)$$

where A and B are complex constants and Y_ν is the Bessel function of the second kind. One may show that A is arbitrary while B depends on λ and n and never vanishes. Consequently the second part of (3.2) is characteristic for the solution in the exceptional cases.

Outside the sphere $|x|=t$ a fundamental solution is given by

$$(3.4)\qquad G_\lambda(t,x)=t^{1-\lambda}\hat{\rho}^{(\lambda-n-1)/2}K_{(n+1-\lambda)/2}(\hat{\rho}c)$$

where $\hat{\rho}=(t^2-x^2)_-^{1/2}$ and K_ν is the modified Bessel function. In the solution (3.2) both the expressions (3.3) and (3.4) will appear.

The fundamental solution displays singularities on the light cone. As in the case where $c=0$ we find that the singularities are of different type depending on n being even or odd. It is also easily seen that classical solutions can be obtained under the same conditions on $f(x)$ as in § 2.4. However, even under these conditions there may occur logarithmic singularities in certain time-derivatives of $u(t,x;\lambda;c)$. This may be seen in the following way (also cf. [3]):

The Fourier transform of $u(t,x;\lambda;c)$ always contains, according to (2.6), a term

$$C\cdot\{(k^2+c^2)^{1/2}t\}^{(1-\lambda)/2}Y_{(1-\lambda)/2}\{(k^2+c^2)^{1/2}t\}\cdot\hat{f}(k).$$

From the behavior of the function $Y_\nu(z)$ near $z=0$ it is easily seen that for $t\to 0$ the derivative of order $1-\lambda$,

$$\left(\frac{\partial}{\partial t}\right)^{1-\lambda}\hat{u}(t,k;\lambda;c),$$

behaves like

$$\{(k^2+c^2)^{1/2}\}^{1-\lambda}Y_0\{(k^2+c^2)^{1/2}t\}\cdot\hat{f}(k).$$

Consequently it has a logarithmic singularity at $t=0$ which vanishes if

$$(k^2+c^2)^{(1-\lambda)/2}\hat{f}(k)=0.$$

As a final result we obtain (in analogy to [3, § 4]) that in the cases where $\lambda=-1,-3,-5,\cdots$, the time derivative of order $1-\lambda$ of a solution has a logarithmic singularity unless we assume that the initial condition $f(x)$ satisfies

$$(\Delta-c^2)^{(1-\lambda)/2}f(x)=0.$$

This generalizes an earlier result of Blum [1].

REFERENCES

[1] E. K. BLUM, *The Euler–Poisson–Darboux equation in the exceptional cases*, Proc. Amer. Math. Soc., 5 (1954), pp. 511–520.

[2] D. W. BRESTERS, *Initial value problems for iterated wave operators*, Thesis, Enschede, Netherlands, 1969.

[3] ———, *On the equation of Euler–Poisson–Darboux*, this Journal, 4 (1973), pp. 31–41.

[4] R. W. CARROLL, *Some singular Cauchy problems*, Ann. Mat. Pura Appl. Ser. IV, 56 (1961), pp. 1–31.

[5] R. W. CARROLL AND R. E. SHOWALTER, *Singular and Degenerate Cauchy Problems*, Academic Press, New York, 1976.

[6] J. M. GEL'FAND AND G. J. SHILOV, *Verallgemeinerte Funktionen*, Bd. I, Deutscher Verlag der Wissenschaften, Berlin, 1960.

[7] E. M. DE JAGER, *Theory of Distributions*, Mathematics Applied to Physics, E. Roubine, ed. Springer-Verlag, New York, 1970, pp. 52–110.

[8] A. OSSICINI, *Problema singolare di Cauchy, relativo ad una generalizzazione dell' equazione di Eulero–Poisson–Darboux*, Atti Accad. Naz. Lincei. Rend. Cl. Sci. Fis. Mat. Natur., 35 (1963), pp. 454–459.

[9] ———, *Problema singolare di Cauchy, relativo ad una generalizzione dell' equazione di Eulero-Poisson–Darboux. Il caso* $k \leqq p-1$, Rend. Mat. e Appl., 23 (1964), pp. 40–65.

[10] A. WEINSTEIN, *On the wave equation and the equation of Euler–Poisson*, Proc. Symposia Appl. Math. Vol. 5, McGraw-Hill, New York, 1954, pp. 137–147.

[11] E. T. YOUNG, *On a generalized E.P.D. equation*, J. Math. Mech., 18 (1969), pp. 1167–1175.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905/0018 $01.00/0

THE SCATTERING OF ACOUSTIC WAVES BY A SPHERICALLY STRATIFIED MEDIUM AND AN OBSTACLE*

DAVID COLTON†

Abstract. The problem of the multiple scattering of a plane acoustic wave by a quasi-homogeneous spherically stratified medium of compact support and a bounded obstacle is considered. It is shown that this problem can be reformulated as a singular integral equation over the boundary of the scattering obstacle alone. The integral equation can be regularized and inverted in the space of continuous functions defined over the boundary of the scattering obstacle.

1. Introduction. In problems associated with the scattering of acoustic waves the method of integral equations is probably the most satisfactory method for obtaining approximations to the solution in the case when various asymptotic methods fail (cf. [14]). This is chiefly due to the fact that the solution is defined over an unbounded domain, thus presenting severe difficulties in the use of other approximation methods such as finite differences. Nevertheless many problems still remain in connection with the method of integral equations, particularly in connection with the stability of approximation methods near an eigenvalue of the integral equation (cf. [2]). In propagation problems in an inhomogeneous medium additional difficulties are presented in the use of the method of integral equations in that the integral operators are now defined over a region in $\mathbb{R}^3$ (cf. [15]), thus making approximate quadrature schemes rather prohibitive. However in those cases where the inhomogeneous medium is of rather simple nature, e.g., when it is spherically stratified and of compact support, it seems reasonable to expect that in many instances the use of volume integrals can be eliminated and the problem reformulated in terms of an integral equation defined only over the boundary of the scattering obstacle. That this is indeed the case was shown in [4] where it was assumed that the scattering obstacle was starlike with respect to the origin. This was accomplished by combining the method of integral equations as recently extended by Ursell [12] and Jones [9] with that of the theory of integral operators in exterior domains as developed by the author and Wolfgang Wendland [3], [4]. In this paper we shall continue this work and consider the case when the scattering obstacle is no longer starlike with respect to the origin and is in fact completely separated from the spherically stratified medium. Although this distinction makes little difference in the case of the integral equation method involving volume integrals, it plays a crucial role in attempting to formulate the problem as an integral equation over the boundary of the scattering obstacle, and a completely new approach must be derived. This we shall do in this paper, basing our approach on R. P. Gilbert's "method of ascent" [7], [8] and the work of Leis [10] and Brakhage and Werner [1] on the reduced wave equation in a homogeneous medium (for recent extensions of the work of Brakhage, Werner and Leis see [11]).

Mathematically we can formulate our problem as follows. Let a plane wave of frequence ω moving in the direction of the z axis be scattered by a quasi-homogeneous spherically stratified medium B of compact support outside of which is situated a "hard" bounded scattering body D with smooth boundary ∂D. Let $c(r)$ ($r=|\mathbf{x}|$ for $\mathbf{x}\in\mathbb{R}^3$) denote the local speed of sound and assume that $c(r)=c_0=a$

* Received by the editors November 4, 1976.

† Department of Mathematics, University of Strathclyde, Glasgow, Scotland. This research was supported in part by the U.S. Air Force Office of Scientific Research under Grant 76-2879.

constant for $r>a$. Then after factoring out a term of the form $e^{-i\omega t}$ we are led to the problem of determining the velocity potential $u(\mathbf{x})$ from the equations

$$u(\mathbf{x})=e^{ikz}+u_s(\mathbf{x}), \tag{1.1}$$

$$\Delta_3 u+k^2B(r)u=0 \quad \text{in } \mathbb{R}^3\backslash\bar{D}, \tag{1.2}$$

$$\frac{\partial u}{\partial \nu}=0 \quad \text{on } \partial D, \tag{1.3}$$

$$\lim_{r\to\infty} r\left(\frac{\partial u_s}{\partial r}-iku_s\right)=0, \tag{1.4}$$

where $k=\omega/c_0>0$, $B(r)=(c_0/c(r))^2$, ν is the inward pointing unit normal to ∂D, and $u_s(\mathbf{x})$ denotes the velocity potential of the scattered wave. We shall make the assumption that $B(r)$ is a continuously differentiable function of r. In § 2 of this paper we shall reformulate (1.1)–(1.4) as a singular integral equation over ∂D, and in § 3 we shall show that a unique solution of this integral equation exists in $C(\partial D)$, the Banach space of continuous functions over ∂D with respect to the maximum norm. Since it shall be seen that the above mentioned singular integral equation can be regularized, a variety of known methods are available for its numerical solution (cf. [5]).

2. The formulation of (1.1)–(1.4) as a singular integral equation over ∂D. We shall look for a solution of (1.1)–(1.4) in the form

$$\begin{aligned} u(\mathbf{x})=&\frac{1}{2\pi}\int_{\partial D}\mu(\boldsymbol{\xi})\left[\frac{e^{ikR}}{R}+\alpha\frac{\partial}{\partial\nu_\xi}\frac{e^{ikR}}{R}\right]d\omega+\sum_{n=0}^{\infty}\sum_{m=-n}^{n}a_{nm}h_n^{(1)}(kr)P_n^m(\cos\theta)\,e^{im\phi}\\ &+\sum_{n=0}^{\infty}(2n+1)i^n j_n(kr)P_n(\cos\theta), \qquad \mathbf{x}\in\mathbb{R}^3\backslash\bar{D}\cup B\\ =&\sum_{n=0}^{\infty}\sum_{m=-n}^{n}b_{nm}u_n(r)P_n^m(\cos\theta)\,e^{im\phi}; \qquad \mathbf{x}\in B \end{aligned} \tag{2.1}$$

where $R=|\mathbf{x}-\boldsymbol{\xi}|$, $\mathbf{x}=(r,\theta,\phi)$ in spherical coordinates, α is a complex constant such that $\operatorname{Im}\alpha\neq 0$, $h_n^{(1)}(kr)$ is a spherical Hankel function, $j_n(kr)$ is a spherical Bessel function, $P_n^m(\cos\theta)$ is an associated Legendre polynomial, $P_n(\cos\theta)$ a Legendre polynomial, and

$$u_n(r)=r^n\left[1-2r\int_0^1\sigma^{2n+2}R_3(r,r;r\sigma^2,0)\,d\sigma\right], \tag{2.2}$$

where $R(x,y;x_0,y_0)$ is the Riemann function for

$$R_{xy}+\frac{k^2}{4}B(\sqrt{xy})R=0 \tag{2.3}$$

with the subscript denoting differentiation with respect to x_0. The density $\mu(\boldsymbol{\xi})$ and the constants a_{nm}, b_{nm}, $n=0,1,2,\cdots$, $-n\leqq m\leqq n$, are unknown and are to be determined from (1.1)–(1.7). The parameter α is introduced in order to insure the invertibility of the integral equation arising from (2.1) (see § 3), the functions $u_n(r)$ are solutions of the differential equation

$$\frac{d^2u_n}{dr^2}+\frac{2}{r}\frac{du_n}{dr}+\left[k^2B(r)-\frac{n(n+1)}{r^2}\right]u_n=0 \tag{2.4}$$

represented by means of R. P. Gilbert's "method of ascent" [7], [8], and we have used Sonine's formula to expand e^{ikz}. Note that if the constants a_{mn} and b_{nm} are such that $u(\mathbf{x})$ as defined by (2.1) is continuously differentiable across ∂B, then one can conclude that in fact $u(\mathbf{x}) \in C^2(R^3 \backslash D)$ and satisfies (1.1), (1.2) and (1.4). Assuming for the time being the convergence of the series in (2.1), we have from (1.3) and the discontinuity properties of single and double layer potentials that

$$\begin{aligned} -\frac{\partial}{\partial \nu_x} e^{ikz} &= \mu(\mathbf{x}) + \frac{1}{2\pi} \int_{\partial D} \mu(\boldsymbol{\xi}) \frac{\partial}{\partial \nu_x} \frac{e^{ikR}}{R}\, d\omega \\ &\quad + \frac{\alpha}{2\pi} \frac{\partial}{\partial \nu_x} \int_{\partial D} \mu(\boldsymbol{\xi}) \frac{\partial}{\partial \nu_\xi} \frac{e^{ikR}}{R}\, d\omega \\ &\quad + \sum_{n=0}^{\infty} \sum_{m=-n}^{n} a_{nm} \frac{\partial}{\partial \nu_x} [h_n^{(1)}(kr) P_n^m(\cos\theta)\, e^{im\phi}], \qquad \mathbf{x} \in \partial D. \end{aligned} \tag{2.5}$$

Note that the differentiation of the second integral with respect to ν_x cannot be taken under the integral sign (In [4] the problem of singular integrals was avoided by using the representation due to Jones [9] and Ursell [12]; however for technical reasons this approach is not suitable for the present problem since D does not contain the origin.)

If we now set $\boldsymbol{\xi} = (\rho, \theta_0, \phi_0)$ in spherical coordinates and require $u(\mathbf{x})$ to be continuously differentiable across ∂B, we have, using the addition formulae

$$\begin{aligned} \frac{e^{ikR}}{R} &= \sum_{n=0}^{\infty} (2n+1) j_n(kr) h_n^{(1)}(k\rho) P_n(\cos\gamma), \qquad r < \rho, \\ \cos\gamma &= \cos\theta\cos\theta_0 + \sin\theta\sin\theta_0\cos(\phi - \phi_0), \\ P_n(\cos\gamma) &= \sum_{m=-n}^{n} \frac{(n-m)!}{(n+m)!} P_n^m(\cos\theta) P_n^m(\cos\theta_0) e^{im(\phi-\phi_0)}, \end{aligned} \tag{2.6}$$

the following algebraic system that must be satisfied by a_{nm} and b_{nm}:

$$\begin{aligned} b_{nm} u_n(a) &= a_{nm} h_n^{(1)}(ka) + \delta_{m0}(2n+1) i^n j_n(ka) \\ &\quad + \frac{\gamma_{nm}}{2\pi} \int_{\partial D} \mu(\boldsymbol{\xi}) h_n^{(1)}(k\rho) P_n^m(\cos\theta_0)\, e^{-im\phi_0}\, d\omega \\ &\quad + \alpha \frac{\gamma_{nm}}{2\pi} \int_{\partial D} \mu(\boldsymbol{\xi}) \frac{\partial}{\partial \nu_\xi} [h_n^{(1)}(k\rho) P_n^m(\cos\theta_0)\, e^{-im\phi_0}]\, d\omega, \\ b_{nm} \frac{du_n(a)}{dr} &= a_{nm} \frac{dh_n^{(1)}(ka)}{dr} + \delta_{m0}(2n+1) i^n \frac{dj_n(ka)}{dr} \\ &\quad + \frac{\gamma_{nm}^*}{2\pi} \int_{\partial D} \mu(\boldsymbol{\xi}) h_n^{(1)}(k\rho) P_n^m(\cos\theta_0)\, e^{-im\phi_0}\, d\omega \\ &\quad + \alpha \frac{\gamma_{nm}^*}{2\pi} \int_{\partial D} \mu(\boldsymbol{\xi}) \frac{\partial}{\partial \nu_\xi} [h_n^{(1)}(k\rho) P_n^m(\cos\theta_0)\, e^{-im\phi_0}]\, d\omega, \end{aligned} \tag{2.7}$$

where δ_{ij} denotes the Kronecker delta and

$$\begin{aligned} \gamma_{nm} &= \frac{(n-m)!}{(n+m)!}(2n+1) j_n(ka), \\ \gamma_{nm}^* &= \frac{(n-m)!}{(n+m)!}(2n+1) \frac{dj_n(ka)}{dr}. \end{aligned} \tag{2.8}$$

The system (2.7) has a unique solution provided

$$A_n = \det \begin{vmatrix} -h_n^{(1)}(ka) & u_n(a) \\ -\dfrac{d}{dr}h_n^{(1)}(ka) & \dfrac{du_n(a)}{dr} \end{vmatrix} \neq 0. \tag{2.9}$$

LEMMA. $A_n \neq 0$ *for any* $n \geqq 0$.

Proof. If $A_n = 0$, then from the theory of ordinary differential equations $u_n(r)$ and $h_n^{(1)}(kr)$ are linearly dependent for $r \geqq a$. But this is impossible since $u_n(r)$ is real and not identically zero.

Using Cramer's rule to solve the system (2.7) we have

$$\begin{aligned} a_{nm} &= \frac{c_n}{A_n}\left[\delta_{m0} + \frac{i^{-n}}{2\pi}\frac{(n-m)!}{(n+m)!}\int_{\partial D}\mu(\boldsymbol{\xi})\left(1+\alpha\frac{\partial}{\partial\nu_\xi}\right)h_n^{(1)}(k\rho)P_n^m(\cos\theta_0)\,e^{-im\phi_0}\,d\omega\right] \\ b_{nm} &= \frac{c_n^*}{A_n}\left[\delta_{m0} + \frac{i^{-n}}{2\pi}\frac{(n-m)!}{(n+m)!}\int_{\partial D}\mu(\boldsymbol{\xi})\left(1+\alpha\frac{\partial}{\partial\nu_\xi}\right)h_n^{(1)}(k\rho)P_n^m(\cos\theta_0)e^{-im\phi_0}\,d\omega\right] \end{aligned} \tag{2.10}$$

where

$$\begin{aligned} c_n &= (2n+1)i^n \det \begin{vmatrix} j_n(ka) & u_n(a) \\ \dfrac{dj_n(ka)}{dr} & \dfrac{du_n(a)}{dr} \end{vmatrix}, \\ c_n^* &= (2n+1)i^n \det \begin{vmatrix} j_n(ka) & h_n^{(1)}(ka) \\ \dfrac{dj_n(ka)}{dr} & \dfrac{dh_n^{(1)}(ka)}{dr} \end{vmatrix} = \frac{i^{n+1}(2n+1)}{ka^2}. \end{aligned} \tag{2.11}$$

Assuming the convergence of the resulting series we now have from (2.10) and (2.5) that $\mu(\mathbf{x})$ is the solution of the integral equation

$$\begin{aligned} f(\mathbf{x}) - \frac{\partial}{\partial\nu_x}e^{ikz} &= (\mathbf{I}+\mathbf{T})\mu \\ &= \mu(\mathbf{x}) + \frac{1}{2\pi}\int_{\partial D}\mu(\boldsymbol{\xi})\frac{\partial}{\partial\nu_x}\left[\frac{e^{ikR}}{R} + \Gamma(\mathbf{x};\boldsymbol{\xi})\right]d\omega \\ &\quad + \frac{\alpha}{2\pi}\frac{\partial}{\partial\nu_x}\int_{\partial D}\mu(\boldsymbol{\xi})\frac{\partial}{\partial\nu_\xi}\frac{e^{ikR}}{R}\,d\omega, \qquad \mathbf{x}\in\partial D, \end{aligned} \tag{2.12}$$

where

$$f(\mathbf{x}) = -\sum_{n=0}^{\infty}\frac{c_n}{A_n}\frac{\partial}{\partial\nu_x}[h_n^{(1)}(kr)P_n(\cos\theta)],$$

$$\begin{aligned} \Gamma(\mathbf{x};\boldsymbol{\xi}) &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n}\frac{c_n i^{-n}(n-m)!}{A_n(n+m)!}\left(1+\alpha\frac{\partial}{\partial\nu_\xi}\right)[h_n^{(1)}(k\rho)P_n^m(\cos\theta)P_n^m(\cos\theta_0)\,e^{im(\phi-\phi_0)}] \\ &= \sum_{n=0}^{\infty}\frac{c_n i^{-n}}{A_n}\left(1+\alpha\frac{\partial}{\partial\nu_\xi}\right)[h_n^{(1)}(k\rho)h_n^{(1)}(kr)P_n(\cos\gamma)]. \end{aligned} \tag{2.13}$$

We shall now show that the series appearing in (2.12) are convergent. From the asymptotic behavior of spherical Bessel and Hankel functions and (using (2.2))

$$u_n(a)=a^n\left(1+O\left(\frac{1}{n}\right)\right),\qquad \frac{du_n(a)}{dr}=na^{n-1}\left(1+O\left(\frac{1}{n}\right)\right), \tag{2.14}$$

we have

$$\left|\frac{c_n}{A_n}\right|\leqq C_1\left(\frac{(ak)^{2n}}{2^{2n}n!(n+1)!}\right),\qquad \left|\frac{c_n^*}{A_n}\right|\leqq C_2\left(\frac{(ak)^n}{2^n n!}\right), \tag{2.15}$$

where the constants C_1 and C_2 are independent of n. From (2.15) and the inequalities (cf. [6])

$$\begin{aligned} &|P_n(\cos\ \gamma)|\leqq 1,\\ &|P_n'(\cos\gamma)|\leqq \tfrac{1}{2}n(n+1),\\ &|h_n^{(1)}(k\rho)|\leqq C_3\frac{n!2^n}{(k\rho)^n},\\ &\left|\frac{dh_n^{(1)}(k\rho)}{d\rho}\right|\leqq C_4\frac{(n+1)!2^n}{(k\rho)^n}, \end{aligned} \tag{2.16}$$

where C_5 and C_6 are constants. Since $\bar{B}\cap\bar{D}=\emptyset$ we have that for $\mathbf{x}$, $\boldsymbol{\xi}\in\partial D$, $\rho>a>0$, $\boldsymbol{\xi}\in\partial D$ that the series appearing in (2.12) are majorized by

$$\begin{aligned} &\left|\frac{\partial}{\partial\nu_x}\Gamma(\mathbf{x};\boldsymbol{\xi})\right|\leqq C_5\sum_{n=0}^{\infty}n^2(n+1)^2\left(\frac{a^2}{\rho r}\right)^n,\\ &|f(\mathbf{x})|\leqq C_6\sum_{n=0}^{\infty}\frac{n+1}{2^n n!}\left(\frac{a^2k}{r}\right)^n \end{aligned} \tag{2.17}$$

where C_5 and C_6 are constants. Since $\bar{B}\cap\bar{D}=\emptyset$ we have that for $\mathbf{x}$, $\boldsymbol{\xi}\in\partial D$, $p>a>0$, $r>a>0$, and hence the series in (2.17) (and hence (2.12)) are convergent. More generally, the above estimates show that the series in (2.13) are absolutely and uniformly convergent for $\mathbf{x}\in\mathbb{R}^3\backslash B$, $\boldsymbol{\xi}\in\partial D$, and hence if $\mu(\mathbf{x})\in C(\partial D)$ can be determined from (2.12), (2.1) gives $u(\mathbf{x})$ for $x\in\mathbb{R}^3\backslash D\cup B$. For $\mathbf{x}\in B$ we have from similar estimates that the series for $u(\mathbf{x})$ in B (as given by (2.1)) is majorized by the series

$$C_7\sum_{n=0}^{\infty}\frac{a^{2n}k^n}{2^n n!}+C_8\sum_{n=0}^{\infty}n(n+1)\left(\frac{a}{\rho}\right)^n, \tag{2.18}$$

where C_7 and C_8 are constants. Since $\rho>a$ it is seen that the series (2.18) is convergent. Hence if $\mu(\mathbf{x})\in C(\partial D)$ can be determined from (2.12), (2.1) and (2.10) give the solution of (1.1)–(1.4). (2.12) is a singular integral equation over ∂D; however it can be regularized (cf. [2], [10], [11]) and hence is amenable to analytic and numerical approximation methods, provided the operator $\mathbf{I}+\mathbf{T}$ is invertible in $C(\partial D)$. Therefore to complete the discussion of our problem (1.1)–(1.4) it remains to be shown that $(\mathbf{I}+\mathbf{T})^{-1}$ exists in $C(\partial D)$. This shall be the topic of the next section.

3. The invertibility of I+T in $C(\partial D)$. In order to establish the invertibility of the operator $\mathbf{I}+\mathbf{T}$ in $C(\partial D)$ it is first necessary to establish the following uniqueness result for the solution to (1.1)–(1.4). The proof of the lemma below proceeds along classical lines (cf. [13]) except for the conclusion where, due to the fact that $B(r)$ is not analytic, it is no longer possible to appeal to the analyticity of solutions to (1.2).

LEMMA. *Let* $u(\mathbf{x})\in C^2(\mathbb{R}^3\backslash\bar{D})\cap C^1(\mathbb{R}^3\backslash D)$ *be a solution of* (1.2) *in the exterior of D satisfying the Sommerfeld radiation condition* (1.4) *at infinity and the boundary condition* $\partial u/\partial\nu=0$ *on* ∂D. *Then* $u(\mathbf{x})=0$ *for* $\mathbf{x}\in\mathbb{R}^3\backslash D$.

Proof. Let Ω be a ball centered at the origin containing $B\cup D$ in its interior. Then from Green's formula

$$\iint_{\Omega\backslash D}(u\Delta\bar{u}-\bar{u}\Delta u)\,dv=-\int_{\partial D}\left(\bar{u}\frac{\partial u}{\partial\nu}-u\frac{\partial\bar{u}}{\partial\nu}\right)d\omega-\int_{\partial\Omega}\left(\bar{u}\frac{\partial u}{\partial r}-u\frac{\partial\bar{u}}{\partial r}\right)d\omega. \tag{3.1}$$

Since k and $B(r)$ are real and $\partial u/\partial\nu=\partial\bar{u}/\partial\nu=0$ on ∂D we have from (3.1) that

$$\int_{\partial\Omega}\left(\bar{u}\frac{\partial u}{\partial r}-u\frac{\partial\bar{u}}{\partial r}\right)d\omega=0. \tag{3.2}$$

But outside $B\cup D$, $u(\mathbf{x})$ is a solution of $\Delta_3u+k^2u=0$ satisfying the Sommerfeld radiation condition and hence in $\mathbb{R}^3\backslash\Omega$ we can expand $u(\mathbf{x})$ in the form

$$u(\mathbf{x})=\sum_{n=0}^{\infty}\sum_{n=-m}^{m}\tilde{a}_{nm}h_n^{(1)}(kr)P_n^m(\cos\theta)\,e^{im\phi}, \tag{3.3}$$

where the series (3.3) converges absolutely and uniformly. By the orthogonality of spherical harmonics and the formula

$$\overline{h_n^{(1)}(kr)}\frac{d}{dr}h_n^{(1)}(kr)-h_n^{(1)}(kr)\frac{d}{dr}\overline{h_n^{(1)}(kr)}=\frac{2i}{kr^2} \tag{3.4}$$

we can conclude that from (3.2) and (3.3) that

$$\sum_{n=0}^{\infty}\sum_{n=-m}^{m}|\tilde{a}_{nm}|^2=0, \tag{3.5}$$

which implies that $u(\mathbf{x})=0$ in $\mathbb{R}^3\backslash\Omega$. By the analyticity of solutions to $\Delta_3u+k^2u=0$ we can conclude that $u(\mathbf{x})\equiv0$ in $\mathbb{R}^3\backslash B\cup D$. But inside $\bar{B}$ we can expand $u(\mathbf{x})$ in the form

$$u(\mathbf{x})=\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\tilde{b}_{nm}u_n(r)P_n^m(\cos\theta)\,e^{im\phi}, \tag{3.6}$$

and from the fact that $u(\mathbf{x})$ is continuously differentiable across $r=a$ we have

$$\tilde{b}_{nm}u_n(a)=\tilde{b}_{nm}\frac{d}{dr}u_n(a)=0. \tag{3.7}$$

If $\tilde{b}_{nm}\neq0$ then $u_n(r)=0$ from the theory of ordinary differential equations since $u_n(r)$ is a solution of (2.4). But $u_n(r)\not\equiv0$ and hence $\tilde{b}_{nm}=0$ for $n\geqq0$, $-n\leqq m\leqq n$ which now implies $u(\mathbf{x})=0$ in $\mathbb{R}^3\backslash D$.

We are now in a position to establish the invertibility of $\mathbf{I}+\mathbf{T}$ in $C(\partial D)$.

THEOREM. *The operator* $(\mathbf{I}+\mathbf{T})^{-1}$ *exists in* $C(\partial D)$. *Hence the solution to* (1.1)–(1.4) *exists and is given by* (2.1), (2.10) *where* $\mu=(\mathbf{I}+\mathbf{T})^{-1}[f(\mathbf{x})-(\partial/\partial\nu_x)\,e^{ikz}]$ *with* $f(\mathbf{x})$ *defined by* (2.15).

Proof. The operator $\mathbf{I}+\mathbf{T}$ is a singular integral operator. However, as previously mentioned, it can be regularized, and hence the Fredholm alternative is valid. Let μ be a solution of $(\mathbf{I}+\mathbf{T})\mu=0$ and define $u(\mathbf{x})$ by

$$u(\mathbf{x})=\begin{cases}\dfrac{1}{2\pi}\displaystyle\int_{\partial D}\mu(\boldsymbol{\xi})\left[\frac{e^{ikR}}{R}+\alpha\frac{\partial}{\partial\nu_\xi}\frac{e^{ikR}}{R}\right]d\omega \\ \qquad +\displaystyle\sum_{n=0}^{\infty}\sum_{m=-n}^{n}a_{nm}h_n^{(1)}(kr)P_n^m(\cos\theta)\,e^{im\phi}, & \mathbf{x}\in\mathbb{R}^3\backslash D\cup B,\\ \displaystyle\sum_{n=0}^{\infty}\sum_{m=-n}^{n}b_{nm}u_n(r)P_n^m(\cos\theta)\,e^{im\phi}, & \mathbf{x}\in B,\end{cases}\tag{3.8}$$

where the constants a_{nm} and b_{nm} are given by (2.10) with the terms $(c_n)/(A_n)\delta_{m0}$ and $(c_n^*)/(A_n)\delta_{m0}$ absent. Then (3.8) defines a solution $u(\mathbf{x})$ of (1.1) satisfying the Sommerfeld radiation condition (1.4) at infinity, and since $(\mathbf{I}+\mathbf{T})\mu=0$ we have $\partial u/\partial\nu=0$ on ∂D. From the preceding lemma we can now conclude that $u(\mathbf{x})=0$ outside D. Let $b>a$ be the distance of D from the origin. Then, using the addition formula for Bessel and Legendre functions as in § 2 of this paper we have for $n\geqq 0,\ -n\leqq m\leqq n$,

$$a_{nm}h_n^{(1)}(kr)+\left[\frac{2n+1}{2\pi}\frac{(n-m)!}{(n-m)!}\int_{\partial D}\mu(\boldsymbol{\xi})\left(1+\alpha\frac{\partial}{\partial\nu_\xi}\right)h_n^{(1)}(k\rho)P_n^m(\cos\theta_0)e^{-im\phi_0}\,d\omega\right]j_n(kr)=0\tag{3.9}$$

for $a<r<b$. Since $h_n^{(1)}(kr)$ and $j_n(kr)$ are linearly independent, (3.9) implies that $a_{nm}=0$ for $n\geqq 0,\ -n\leqq m\leqq n$. Since $u(\mathbf{x})=0$ outside D we now have from (3.8) that

$$0=-\alpha\mu(\mathbf{x})+\frac{1}{2\pi}\int_{\partial D}\mu(\boldsymbol{\xi})\frac{e^{ikR}}{R}d\omega+\frac{\alpha}{2\pi}\int_{\partial D}\mu(\boldsymbol{\xi})\frac{\partial}{\partial\nu_\xi}\frac{e^{ikR}}{R}d\omega,\qquad \mathbf{x}\in\partial D.\tag{3.10}$$

But from (3.10) and [1] we can conclude that $\mu(\mathbf{x})=0$ for $\mathbf{x}\in\partial D$, and hence by the Fredholm alternative $(\mathbf{I}+\mathbf{T})^{-1}$ exists in $C(\partial D)$.

We note in closing that the operator $\mathbf{I}+\mathbf{T}$ can be regularized (cf. [2], [10], [11]) to yield an equation of the form $(\mathbf{I}+\mathbf{K})\mu=g$ where $\mathbf{K}$ is a compact operator defined on $C(\partial D)$, and hence a wide variety of procedures are available for approximating the density $\mu(\mathbf{x})$ (cf. [5]).

REFERENCES

[1] H. Brakhage and P. Werner, *Uber das Dirichletsche Aussenraumproblem fur die Helmholtzche Schwingungsgleichung*, Arch. Math. 16 (1965), pp. 325–329.

[2] A. J. Burton, *Numerical solution of scalar diffraction problems*, Numerical Solution of Integral Equations, L. M. Delves and J. Walsh, eds., Clarendon Press, Oxford, England, 1974.

[3] D. Colton, *Solution of Boundary Value Problems by the Method of Integral Operators*, Pitman Press, London, 1976.

[4] D. Colton and W. Wendland, *Constructive methods for solving the exterior Neumann problem for the reduced wave equation in a spherically symmetric medium*, Proc. Royal Soc. Edinburgh Sect. A 75 (1976), pp. 97–107.

[5] L. M. Delves and J. Walsh, *Numerical Solution of Integral Equations*, Clarendon Press, Oxford, England, 1974.

[6] A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, *Higher Trancendental Functions*, vol. II, McGraw-Hill, New York, 1953.

[7] R. P. Gilbert, *Constructive Methods for Elliptic Equations*, Springer-Verlag, Berlin, 1974.

[8] R. P. Gilbert and P. Linz, *The numerical solution of some elliptic boundary value problems by integral operator methods*, Constructive and Computational Methods for Differential and Integral Equations, D. L. Colton and R. P. Gilbert, eds., Springer-Verlag, Berlin, 1974.

[9] D. S. JONES, *Integral equations for the exterior acoustic problem*, Quart. J. Mech. Appl. Math., 27 (1974), pp. 129–142.

[10] R. LEIS, *Vorlesungen über Partielle Differentialgleichungen Zweiter Ordung*, BI Mannheim, Mannheim, 1967.

[11] G. F. ROACH AND R. KRESS, *On mixed boundary value problems for the Helmholtz equation*, Proc. Royal Soc. Edinburgh Sect. A, 77 (1977), pp. 65–77.

[12] F. URSELL, *On the exterior problems of acoustics*, Proc. Cambridge Philos. Soc. 74 (1973), pp. 117–125.

[13] I. N. VEKUA, *New Methods for Solving Elliptic Equations*, North-Holland, Amsterdam, 1967.

[14] R. WAIT, *Use of finite elements in multidimensional problems in practice*, Numerical Solution of Integral Equations, L. M. Delves and J. Walsh, eds., Clarendon Press, Oxford, England, 1974.

[15] P. WERNER, *Randwertproblem der mathematischen Akustik*, Arch. Rational Mech. Anal. 10 (1962), pp. 29–66.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0019 $01.00/0

COMPARISON THEOREMS FOR LINEAR BOUNDARY VALUE PROBLEMS*

JAMES S. MULDOWNEY†

Abstract. Conditions on a linear differential operator L are given which guarantee the nonexistence of nontrivial solutions to certain homogeneous boundary value problems. Less restrictive conditions for the nonexistence of nontrivial nonnegative solutions are found and applied to questions of disconjugacy. Some new proofs of known results as well as new disconjugacy criteria are obtained.

1. Introduction. Let L denote the nth order linear differential operator defined by

$$Ly = y^{(n)} + a_1 y^{(n-1)} + \cdots + a_n y,$$

if $y \in AC^{n-1}[a, b]$, where a_k are real valued functions of class C^{n-k} on $[a, b]$. Also let the boundary form $U: AC^{n-1}[a, b] \to R^n$ be defined by

$$U_i y = \sum_{j=1}^{n} (M_{ij} y^{(j-1)}(a) + N_{ij} y^{(j-1)}(b)), \qquad i = 1, \cdots, n$$

where M_{ij}, N_{ij} are real numbers such that the $n \times 2n$ matrix $[M: N]$ has rank n; the form U is then said to have rank n. This paper is concerned with boundary value problems of the form

$$\mathscr{P}(L; U): Lx = f, \qquad Ux = \gamma, \qquad x \in AC^{n-1}[a, b]$$

where $f \in \mathscr{L}_1[a, b]$ and $\gamma \in R^n$. Conditions on the homogeneous problem

$$\mathscr{P}_0(L; U): Lx = 0, \qquad Ux = 0, \qquad x \in C^n[a, b]$$

which guarantee the nonexistence of nontrivial solutions to this problem, and more particularly conditions which ensure the nonexistence of nontrivial nonnegative solutions, are found. In the special case

$$U_i y = y^{(i-1)}(a), \quad i = 1, \cdots, k, \qquad U_i y = y^{(i-k-1)}(b), \quad i = k+1, \cdots, n,$$

the problems $\mathscr{P}(L; U)$, $\mathscr{P}_0(L; U)$ are denoted $\mathscr{P}(L; k, a, b)$ and $\mathscr{P}_0(L; k, a, b)$ respectively. Discussions of boundary value problems may be found in the books of Coddington and Levinson [2] and Reid [14].

The main tools in this paper are the Lagrange identity and the boundary-form formula described below. A complete discussion may be found in [2, Chapter 11]. The Lagrange adjoint L^* of L is defined by

$$L^* y = (-1)^n y^{(n)} + (-1)^{n-1}(a_1 y)^{(n-1)} + \cdots + a_n y$$

and Lagrange's identity is

$$vLu - uL^*v = [uv]', \tag{1.1}$$

if u and v are n times differentiable functions, where

$$[uv] = \sum_{m=1}^{n} \sum_{j+k=m-1} (-1)^j u^{(k)} (a_{n-m} v)^{(j)}$$

* Received by the editors September 2, 1976 and in revised form March 4, 1977.

† Department of Mathematics, University of Alberta, Edmonton, Alberta, Canada T6G 2G1. This research was supported by NRC Grant A-7197.

and $a_0 = 1$. For a boundary form U of rank n, there exist boundary forms U_c such that the $2n \times 2n$ matrix

$$\begin{bmatrix} M & N \\ M_c & N_c \end{bmatrix}$$

has rank $2n$; U and U_c are called complementary boundary forms. For each pair of complementary boundary forms U, U_c of rank n, there exists a unique pair of complementary forms U^*, U_c^* such that

$$[xy](b)-[xy](a) = Ux \cdot U_c^* y + U_c x \cdot U^* y \tag{1.2}$$

if the functions x, y are $n-1$ times differentiable at a and b. This is the boundary-form formula. Note that, in general, if U is specified then adjoint boundary forms U^* depend on the coefficients in L, and conversely if U^* is specified U depends on L (cf. [2, Theorem 3.1, p. 289]). However, in the case of the boundary form $U = (k, a, b)$ it is always possible to choose $U^* = (n-k, a, b)$.

The first comparison result proved in this paper, Theorem 3.1, states that if $\mathcal{P}(L^*; U^*)$ has a solution with $\gamma = 0, f \geqq 0$ ($\neq 0$) then there is no nontrivial nonnegative solution to $\mathcal{P}_0(L; U)$. This theorem is extended in Theorems 3.2 and 3.3 by the introduction of a class $\mathcal{H}(U^*)$ of functions $H(t, s)$. It is shown that if $[L_s^* H(\cdot, s)]_-$ is 'small' for some $H \in \mathcal{H}(U^*)$, then $\mathcal{P}_0(L; U)$ has no nontrivial nonnegative solution, and if $|L_s^* H(\cdot, s)|$ is 'small', then $\mathcal{P}_0(L; U)$ has no nontrivial solution.

2. The classes $\mathcal{H}(U^*)$, $\mathcal{N}_p(U)$.

DEFINITION 2.1. A function $H: [a, b] \times [a, b] \to R$ belongs to the class $\mathcal{H}(U^*)$ if, for each $t \in [a, b]$,

(i) $H(t, \cdot) \in C^n([a, t) \cup (t, b]) \cap C^{n-2}[a, b]$,
(ii) $(\partial^{n-1}/\partial s^{n-1})H(t, t+) - (\partial^{n-1}/\partial s^{n-1})H(t, t-) = (-1)^n$,
(iii) $U^* H(t, \cdot) = 0$.

Note that if, in addition to (i), (ii), (iii), $y = H(t, \cdot)$ satisfies $L^* y = 0$ on $[a, t) \cup (t, b]$ for each $t \in [a, b]$ then H is the Green's function for $\mathcal{P}(L; U)$. However, this condition is not required here; indeed if G is any Green's function corresponding to any differential operator L_0 of order n and satisfies $U^* G(t, \cdot) = 0$, then $G \in \mathcal{H}(U^*)$. However $\mathcal{H}(U^*)$ does not consist entirely of Green's functions or generalized Green's functions since, if $H \in \mathcal{H}(U^*)$ and $K: [a, b] \times [a, b] \to R$ is such that

$$K(t, \cdot) \in C^n[a, b], \qquad U^* K(t, \cdot) = 0$$

for each $t \in [a, b]$, then $H + K \in \mathcal{H}(U^*)$.

DEFINITION 2.2. $\mathcal{N}_p(U)$, $1 \leqq p \leqq \infty$, denotes the set of functions $\eta: [a, b] \to R$ such that $\eta \geqq 0$, and $\eta x \in \mathcal{L}_p[a, b]$ if $x \in AC^{n-1}[a, b]$ and $Ux = 0$, where $\mathcal{L}_p[a, b]$ denotes the usual Lebesgue class.

DEFINITION 2.3. If $u_1, \cdots, u_n \in C^n[a, b]$, $r_0, \cdots, r_m$ are nonnegative integers such that $r_0 + \cdots + r_m = n$ and $t_0, \cdots, t_m \in [a, b]$, then

$$\mathcal{W}_n(u_1, \cdots, u_n)\begin{bmatrix} t_0, \cdots, t_m \\ r_0, \cdots, r_m \end{bmatrix}$$

denotes the determinant of the $n \times n$ matrix the kth row of which is

$$[u_k(t_0), \cdots, u_k^{(r_0-1)}(t_0), \cdots, u_k(t_m), \cdots, u_k^{(r_m-1)}(t_m)].$$

In particular $\mathcal{W}_n(u_1, \cdots, u_n)[{}^t_n]$ is the Wronskian determinant $W(u_1, \cdots, u_n)(t) = \det[u_i^{(j-1)}(t)]$.

Throughout this paper, $a_+ = \frac{1}{2}(|a|+a)$ and $a_- = \frac{1}{2}(|a|-a)$ for any real number a.

The following proposition follows from Lagrange's identity (1.1) and the boundary-form formula (1.2). The suffix s in L_s denotes the variable to which the differentiation pertains.

PROPOSITION 2.1. *If $H \in \mathscr{H}(U^*)$ and $x \in AC^{n-1}[a, b]$, then*

$$x(t) = -Ux \cdot U_c^* H(t, \cdot) + \int_a^b H(t, s) Lx(s)\, ds - \int_a^b x(s) L_s^* H(t, s)\, ds. \tag{2.1}$$

Proof. From (1.1) with $u = x$, $v = H(t, \cdot)$, it follows that

$$\begin{aligned} &\int_a^b H(t, s) Lx(s)\, ds - \int_a^b x(s) L_s^* H(t, s)\, ds \\ &\quad = [xH(t, \cdot)](b) - [xH(t, \cdot)](t+) + [xH(t, \cdot)](t-) - [xH(t, \cdot)](a). \end{aligned} \tag{2.2}$$

But from the conditions (i), (ii) of Definition 2.1

$$\begin{aligned} &[xH(t, \cdot)](t-) - [xH(t, \cdot)](t+) \\ &\quad = (-1)^n x(t) \left[\frac{\partial^{n-1}}{\partial s^{n-1}} H(t, t+) - \frac{\partial^{n-1}}{\partial s^{n-1}} H(t, t-) \right] = x(t), \end{aligned} \tag{2.3}$$

and from the boundary-form formula (1.2)

$$\begin{aligned} [xH(t, \cdot)](b) - [xH(t, \cdot)](a) &= Ux \cdot U_c^* H(t, \cdot) + U_c x \cdot U^* H(t, \cdot) \\ &= Ux \cdot U_c^* H(t, \cdot), \end{aligned} \tag{2.4}$$

from condition (iii) of Definition 2.1. Combining (2.2), (2.3) and (2.4) gives (2.1).

Proposition 2.1 shows that any solution of the problem $\mathscr{P}(L; U)$ must satisfy the integral equation

$$x(t) = -\gamma \cdot U_c^* H(t, \cdot) + \int_a^b H(t, s) f(s)\, ds - \int_a^b x(s) L_s^* H(t, s)\, ds \tag{2.5}$$

if $H \in \mathscr{H}(U^*)$. Proposition 2.2 gives conditions under which a solution of (2.5) is also a solution of $\mathscr{P}(L; U)$.

PROPOSITION 2.2. (a) *If $H \in \mathscr{H}(U^*)$ is such that $\delta \in R^n$, $\Delta \in \mathscr{L}_1[a, b]$ and*

$$0 = -\delta \cdot U_c^* H(t, \cdot) + \int_a^b H(t, s) \Delta(s)\, ds \tag{2.6}$$

for each $t \in [a, b]$ implies $\delta = 0$ and $\Delta = 0$, then any solution of (2.5) *such that $x \in$ AC$^{n-1}[a, b]$ is a solution of $\mathscr{P}(L; U)$.*

(b) *If $H \in \mathscr{H}(U^*)$ is such that the only solution of*

$$x(t) = -\int_a^b x(s) L_s^* H(t, s)\, ds \tag{2.7}$$

is the zero solution, then a solution of (2.5) *is also a solution of $\mathscr{P}(L; U)$.*

Proof of (a). If x satisfies (2.5) and $x \in$ AC$^{n-1}[a, b]$, then

$$\delta = Ux - \gamma, \qquad \Delta = Lx - f,$$

from Proposition 2.1, satisfy (2.6) so that x satisfies $\mathscr{P}(L; U)$.

The conditions of part (a) are satisfied if $H \in \mathscr{H}(U^*)$ is a Green's function for a boundary value problem associated with any differential operator L_0 of order n and the components of the vector valued function $U_c^*H(t,\cdot)$ are linearly independent on $[a,b]$. In that case the right-hand side of (2.6) satisfies $L_0x=\Delta$, but since $x=0$ it follows that $\Delta=0$; therefore $\delta\cdot U_c^*H(t,\cdot)=0$ and hence $\delta=0$.

Proof of (b). From Proposition 2.1, any solution of $\mathcal{P}_0(L;U)$ satisfies (2.7); thus $\mathcal{P}_0(L;U)$ has only the zero solution. Therefore, given γ and f, $\mathcal{P}(L;U)$ has a unique solution. The conditions of part (b) also imply that (2.5) has a unique solution which must therefore be the unique solution of $\mathcal{P}(L;U)$.

PROPOSITION 2.3. *For each boundary form U of rank n the class $\mathcal{H}(U)$ is nonempty. In particular $\mathcal{H}(U)$ contains a Green's function.*

Proof. If $\phi_i\in C^n[a,b]$, $i=1,\cdots,n$, are such that $W(\phi_1,\cdots,\phi_n)=\det[\phi_i^{(j-1)}]\neq 0$, $\det[U_i\phi_j]\neq 0$ and L_0 is defined by $L_0y=W(\phi_1,\cdots,\phi_n,y)/W(\phi_1,\cdots,\phi_n)$ then $\mathcal{P}_0(L_0;U)$ has only the trivial solutions so $\mathcal{P}(L_0;U)$ has a Green's function $G(t,s)$ and $H\in\mathcal{H}(U)$ if $H(t,s)=(-1)^nG(s,t)$. It will be shown that the constants $\lambda_1,\cdots,\lambda_n$ may be chosen so that $\phi_i(t)=e^{\lambda_i t}$, $i=1,\cdots,n$, satisfy the required conditions. Without loss of generality it may be assumed that $a=0$, $b=1$. If the constants $\lambda_1,\cdots,\lambda_n$ are distinct then $W(\phi_1,\cdots,\phi_n)\neq 0$. Also $U_i\phi_j=\mu_i(\lambda_j)$ where

$$\mu_i(\lambda)=\sum_{k=1}^{n}[M_{ik}\lambda^{k-1}+N_{ik}\lambda^{k-1}e^{\lambda}].$$

An inductive proof is given that the constants $\lambda_1,\cdots,\lambda_n$ may be chosen so that the determinants $\det[\mu_i(\lambda_j)]\neq 0$, $i,j=1,\cdots,m$, $1\leqq m\leqq n$. Choose λ_1 so that $\mu_1(\lambda_1)\neq 0$; this is possible since the numbers M_{1k}, N_{1k} are not all zero. Suppose that $\lambda_1,\cdots,\lambda_{m-1}$, $1<m\leqq n$ have been found so that $\det[\mu_i(\lambda_j)]\neq 0$, $i,j=1,\cdots,m-1$, and that $\det[\mu_i(\lambda_j)]=0$, $i,j=1,\cdots,m$ for all choices of λ_m. It follows that for each choice of λ_m the mth row of the matrix $[\mu_i(\lambda_j)]$ may be expressed as a linear combination of the first $m-1$ rows, i.e.

$$\mu_m(\lambda_j)=\sum_{i=1}^{m-1}c_i\mu_i(\lambda_j),\qquad j=1,\cdots,m.$$

Considering the first $m-1$ of these equations it follows that since $\det[\mu_i(\lambda_j)]\neq 0$, $i,j=1,\cdots,m-1$, the numbers c_i are uniquely determined and, in particular, are independent of λ_m. Thus the mth equation $\mu_m(\lambda_m)=\sum_{i=1}^{m-1}c_i\mu_i(\lambda_m)$, for all λ_m, implies

$$M_{mj}=\sum_{i=1}^{m-1}c_iM_{ij},\qquad N_{mj}=\sum_{i=1}^{m-1}c_iN_{ij},\qquad j=1,\cdots,n,$$

contradicting rank $[M:N]=n$. This contradiction completes the proof.

3. Results. Theorem 3.1 which has a very simple proof may be considered a generalization of the Sturm comparison principle for $\mathcal{P}_0(L;1,a,b)$ when L is a second order differential operator. A number of important results which have other proofs in the literature are corollaries to Theorem 3.1.

THEOREM 3.1. *If there exists $\psi\in \mathrm{AC}^{n-1}[a,b]$ such that*

$$U^*\psi=0,\qquad L^*\psi\geqq 0,\tag{3.1}$$

with strict inequality holding on a set of positive measure, then there is no nontrivial solution x of $\mathcal{P}_0(L;U)$ such that $x\geqq 0$ on $[a,b]$.

Proof. If such a solution x exists, then $x>0$ a.e. on $[a,b]$ since otherwise there would exist a point $\tau\in[a,b]$ such that $x(\tau)=x'(\tau)=\cdots=x^{(n-1)}(\tau)=0$ implying $x=0$

on $[a, b]$. From (1.1) and (1.2)

$$0=[x\psi](b)-[x\psi](a)$$
$$=\int_a^b(\psi Lx-xL^*\psi)=-\int_a^b xL^*\psi<0,$$

and this contradiction proves the result.

A stronger version of Theorem 3.1 in which it need not be assumed that $L^*\psi\geqq 0$ follows from Theorem 3.2 (cf. Corollary 3.2.1).

The operator L is said to be *disconjugate* on an interval I if the only solution of $Lx=0$ having n zeros or more in I counting multiplicities is the trivial solution. A result of Levin [7] and Sherman [15] shows that L is disconjugate on I if and only if the problem $\mathcal{P}_0(L; k, a, b)$ has no nontrivial nonnegative solution for each $[a, b]\subset I$ and $k=1,\cdots,n-1$. Also L is disconjugate on I if and only if L^* is disconjugate on I (cf. Coppel [3, Chap. 3]). These observations will be used to give examples of applications for Theorem 3.1 in the following corollaries.

COROLLARY 3.1.1. *If L is disconjugate on I and $(-1)^{n-k}q_n\geqq 0$, then the problem $\mathcal{P}_0(L+q_n; k, a, b)$ has no nontrivial nonnegative solution when $[a, b]\subset I$.*

Proof. If L is disconjugate on I, then so also is L^* and the solution of $L^*\psi=1$, $\psi(a)=\cdots=\psi^{(n-k-1)}(a)=\psi(b)=\cdots=\psi^{(k-1)}(b)=0$ satisfies $(-1)^{n-k}\psi>0$ on (a, b) if $[a, b]\subset I$. This follows from a theorem of Pólya [13, Thm. V] or from the sign of the appropriate Green's function (cf. [3]). Therefore $(L+q_n)^*\psi=L^*\psi+q_n\psi\geqq L^*\psi>0$ on (a, b).

The following comparison principle was first stated by Levin [7]. The first published proof is due to Nehari [11].

COROLLARY 3.1.2 (Levin, Nehari). *Suppose that*

$$L_iy=Ly+q_{n,i}y,\qquad i=1,2,$$

where $q_{n,1}\leqq 0\leqq q_{n,2}$ on an interval I. If L_1 and L_2 are disconjugate on I, then so also is L.

This follows immediately from Corollary 3.1.1 and the result of Levin and Sherman [7], [15]. The following more general disconjugacy criterion also follows from Theorem 3.1 and the result of Levin and Sherman.

COROLLARY 3.1.3. *A sufficient condition for the disconjugacy of L on an interval I is that, for each $[a, b]\subset I$ and $k=1,\cdots,n-1$, there exists a function $\psi_k\in \mathrm{AC}^{n-1}[a, b]$ such that*

$$\psi_k(a)=\cdots=\psi_k^{(k-1)}(a)=\psi_k(b)=\cdots=\psi_k^{(n-k-1)}(b)=0 \tag{3.2}$$

and $L\psi_k\geqq 0$ on $[a, b]$ with strict inequality on a set of positive measure. A necessary condition is that all such functions should also satisfy $(-1)^{n-k}\psi_k>0$ on (a, b).

The necessity of this condition follows from Pólya's Theorem V [13] or from the sign of the Green's function for $\mathcal{P}(L; k, a, b)$.

COROLLARY 3.1.4. *Suppose $n>2$, $a_1(t)=\cdots=a_{n-1}(t)=0$ and*

$$\rho_n(t)\leqq a_n(t)\leqq\sigma_n(t),\qquad t\in(a, b),$$

where

$$\rho_n(t)=\begin{cases}-\dfrac{1}{\psi_2(t)}\,(n\ \textit{even}),\quad -\dfrac{1}{\psi_1(t)}\,(n\ \textit{odd}), & t\in\left(a,\dfrac{a+b}{2}\right],\\[2ex] -\dfrac{1}{\psi_{n-2}(t)}, & t\in\left[\dfrac{a+b}{2}, b\right),\end{cases}$$

$$\sigma_n(t)=\begin{cases}\dfrac{1}{\psi_1(t)}(n\ even), & \dfrac{1}{\psi_2(t)}(n\ odd), \quad t\in\left(a,\dfrac{a+b}{2}\right],\\ \dfrac{1}{\psi_{n-1}(t)}, & t\in\left[\dfrac{a+b}{2},b\right),\end{cases}$$

and $\psi_k(t)=\psi_k(a,b;t)=(t-a)^k(b-t)^{n-k}/n!$, $k=1,\cdots,n-1$. *Then* L *is disconjugate on* $[a,b]$.

Proof. First observe that $(-1)^{n-k}L\psi_k=\psi_k[1/\psi_k+(-1)^{n-k}a_n]$; thus $(-1)^{n-k}L\psi_k\geqq 0$ if $(-1)^{n-k}a_n\geqq -1/\psi_k$ and $(-1)^{n-k}L\psi_k>0$ on a set of positive measure since $a_n\in C[a,b]$ and $1/\psi_k$ is unbounded on (a,b). Since

$$\rho_n(t)=\max\left\{\frac{-1}{\psi_k(t)}: n-k \text{ even}\right\},\qquad \sigma_n(t)=\min\left\{\frac{1}{\psi_k(t)}: n-k\text{ odd}\right\}$$

decrease and increase respectively if $[a,b]$ is replaced by $[a',b']\subset[a,b]$, it follows that $(-1)^{n-k}L\psi_k(a',b';t)\geqq 0$ for each subinterval $[a',b']$ of $[a,b]$ and $k=1,\cdots,n-1$. Thus L is disconjugate on $[a,b]$, from Corollary 3.1.3.

In the case $n=2$ the conditions of Corollary 3.1.4 reduce to $a_2(t)\leqq\sigma_2(t)=2/((t-a)(b-t))$, $a<t<b$; thus the result in this case neither implies nor is it implied by Lyapunov's inequality (cf. [6, Thm. 5.1, p. 345]). If the functions ψ_k of Corollary 3.1.4 are considered for general operators L then a disconjugacy criterion of Bessmertnyh and Levin [1] is obtained.

COROLLARY 3.1.5 (Bessmertnyh and Levin). *Suppose* $n>2$ *and for each* $t\in[a,b]$

$$\sum_{j=1}^{n-1}\frac{n-j}{j!n}|a_j(t)|(b-a)^j+\frac{(n-1)^{n-1}}{n!n^n}a_n(t)_+(b-a)^n\leqq 1,$$
$$\sum_{j=1}^{n-1}\frac{n-j}{j!n}|a_j(t)|(b-a)^j+\frac{(n-1)^{n-1}}{n!n^n}a_n(t)_-(b-a)^n\leqq 1. \tag{3.3}$$

Then L *is disconjugate on* $[a,b]$.

Remark. The second inequality may be improved as indicated below.

Proof. Let $\psi_{n,k}(t)=\psi_{n,k}(a,b;t)=(t-a)^k(b-t)^{n-k}/n!$ $k=1,\cdots,n-1$. It is asserted that, if $t\in[a,b]$,

$$|\psi_{n,k}(t)|\leqq\frac{(n-1)^{n-1}}{n!n^n}(b-a)^n,\qquad |\psi_{n,k}^{(n-j)}(t)|\leqq\frac{n-j}{j!n}(b-a)^j, \tag{3.4}$$

$j=0,\cdots,n-1$ with strict inequality holding almost everywhere if $j>0$. The result now follows from Corollary 3.1.3 and the observation that both expressions on the left in (3.3) are increasing functions of $b-a$. To prove the assertion (3.4), first observe that the set of all polynomials $\psi(t)$ with $|\psi^{(n)}(t)|=1$, having n zeros in $[a,b]$ and at least one zero at each of the points a, b satisfy $|\psi(t)|\leqq(n-1)^{n-1}/(n!n^n)(b-a)^n$, $t\in[a,b]$, and that equality can be achieved only for $\psi_{n,k}(t)$, $k=1$, $n-1$ and then at a single point only. Now consider the functions $\psi'_{n,k}(t)$. By the preceding observation

$$|\psi'_{n,k}(t)|<\frac{(n-2)^{n-2}}{(n-1)!(n-1)^{n-1}}(b-a)^{n-1},\qquad k=2,\cdots,n-2,$$

and it may be verified directly that

$$|\psi'_{n,k}(t)|\leqq\frac{1}{(n-1)!n}(b-a)^{n-1},\qquad k=1,n-1,$$

with equality holding at an endpoint a, b in each case. Since $((n-2)/(n-1))^{n-2} \leqq \frac{1}{2}$ it follows that

$$\frac{(n-2)^{n-2}}{(n-1)!(n-1)^{n-1}} \leqq \frac{1}{(n-1)!2(n-1)} \leqq \frac{1}{(n-1)!n},$$

and hence

$$|\psi'_{n,k}(t)| \leqq \frac{1}{(n-1)!n}(b-a)^{n-1}, \qquad k=1, \cdots, n-1.$$

To establish the bounds on the higher derivatives observe that $\psi_{n,k}(t) = ((t-a)/n)\psi_{n-1,k-1}(t)$ $k=2, \cdots, n-1$ and proceed by induction on n and j. The omission of the case $k=1$ is unimportant from considerations of symmetry. Since $\psi_{n,k}^{(n)} = (-1)^{n-k}$ and

$$\psi_{n,k}^{(n-j)} = \frac{t-a}{n}\psi_{n-1,k-1}^{(n-j)} + \frac{n-j}{n}\psi_{n-1,k-1}^{(n-j-1)},$$

it follows that if (3.4) holds for $\psi_{n-1,k-1}$, then

$$\begin{aligned} |\psi_{n,k}^{(n-j)}| &\leqq \frac{b-a}{n}\,\frac{(n-j)}{(j-1)!(n-1)}(b-a)^{j-1} + \frac{n-j}{n}\,\frac{(n-j-1)}{j!(n-1)}(b-a)^j \\ &= \frac{n-j}{j!n}(b-a)^j. \end{aligned}$$

Equality is achieved at an endpoint if $k=1$, $n-1$. Therefore, since (3.4) holds when $n=2$, it holds for all n.

The reader will observe that the first inequality in (3.3) is used to infer that $(-1)^{n-k}L\psi_{n,k} \geqq 0$ when $n-k$ is odd and the second pertains to the case when $n-k$ is even. Thus some improvement is possible in the case of the second inequality in (3.3) when n is even since suitable coefficients for $|a_j(t)|$ in that expression may be obtained by maximizing $|\psi_{n,k}^{(n-j)}(t)|$ for even k only. Thus, for example, when n is even the coefficient of $a_n(t)_-$ may be replaced by the smaller number

$$\frac{4(n-2)^{n-2}}{((n/2)!)^2 n^n}(b-a)^n = \max\{|\psi_{n,k}(t)| : k=2, 4, \cdots, n-2\},$$

and the coefficient of $|a_{n-1}(t)|$ may be replaced by

$$\frac{(n-2)^{n-2}}{(n-1)!(n-1)^{n-1}}(b-a)^{n-1} > \max\{|\psi'_{n,k}(t)| : k=2, 4, \cdots, n-2\}.$$

Corollary 3.1.6 is a disconjugacy criterion of a type introduced for higher order equations by Hartman and Levin [8], [4], [5]. It extends to higher order operators the form of the Sturm comparison theorem which states that a second order operator L is disconjugate on $[a, b]$ if there exists $u \in C^2[a, b]$ such that $u > 0$ and $Lu \leqq 0$ on $[a, b)$. It is more restrictive than the best result of Hartman [5] but has a very simple proof. A system of functions $(u_1, \cdots, u_n)$ is called a *Markov system* on an interval I if $u_j \in C^n(I)$ and $W(u_1, \cdots, u_j) > 0$, $j=1, \cdots, n$. Also the symbol $(u_1, \cdots, \hat{u}_j, \cdots, u_n)$ denotes the system $(u_1, \cdots, u_{j-1}, u_{j+1}, \cdots, u_n)$.

COROLLARY 3.1.6 (Hartman, Levin). *Let $[q, b]$ be a compact interval. Suppose there exist functions $u_1, \cdots, u_{n-1} \in C^n[a, b)$ such that*

(i) $(-1)^{n-j}Lu_j \geqq 0$, $j=1, \cdots, n-1$ *on* $[a, b)$

(ii) $(u_1, \cdots, u_{n-1})$ *and* $(u_1, \cdots, \hat{u}_j, \cdots, u_{n-1})$ *are Markov systems on* $[a, b)$, $1 \leqq j < n-1$.

Then L is disconjugate on $[a, b]$.

Proof. If $[a', b'] \subset [a, b)$ choose λ so that if $u_0(t) = e^{-\lambda t}$, $u_n(t) = e^{\lambda t}$ then $(u_0, \cdots, \hat{u}_j, \cdots, u_n)$ are Markov systems on $[a', b']$ and $(-1)^{n-j} L u_j \geqq 0$, $j = 0, \cdots, n$ with strict inequality if $j = 0, n$. Let

$$\psi_k(t) = \psi_k(a', b'; t) = \mathscr{W}_{n+1}(u_0, \cdots, u_n)\begin{bmatrix} a', & b', & t \\ k, & n-k, & 1 \end{bmatrix}$$

so that ψ_k satisfies (3.2) and

$$L\psi_k(t) = \sum_{j=0}^{n} \mathscr{W}_n(u_0, \cdots, \hat{u}_j, \cdots, u_n)\begin{bmatrix} a', & b' \\ k, & n-k \end{bmatrix}(-1)^{n-j} L u_j(t) > 0,$$

$a' \leqq t \leqq b'$, since the determinants $\mathscr{W}_n(u_0, \cdots, \hat{u}_j, \cdots, u_n)\begin{bmatrix} a', & b' \\ k, & n-k \end{bmatrix}$ are all positive by Pólya's Theorem V [13]. It now follows from Corollary 3.1.3 that L is disconjugate on $[a, b)$. To see that b is not the first conjugate point of a observe that the functions $(u_1, \cdots, u_{n-1})$ may be extended to an interval $[c, b)$, $c < a$, and still satisfy conditions (i), (ii) on $[c, b)$. Thus L is disconjugate on $[c, b)$ and hence, by Theorem 7 of Sherman [15], L is disconjugate on $[a, b]$.

THEOREM 3.2. *Suppose* $H \in \mathscr{H}(U^*)$ *and* $\eta \in \mathscr{N}_p(U)$ *exist such that*

$$\int_a^b \left\{ \int_a^b [L_s^* H(t, s)\eta(t)/\eta(s)]_-^q \, ds \right\}^{p/q} dt < 1, \tag{3.5}$$

where $1 \leqq p \leqq \infty$, $1/p + 1/q = 1$. *Then there is no nontrivial solution x of* $\mathscr{P}_0(L; U)$ *such that* $x \geqq 0$ *on* $[a, b]$. *If p or q* $= \infty$, *the corresponding integral in* (3.5) *should be replaced by the appropriate supremum.*

Proof. Suppose x is a nontrivial solution of $\mathscr{P}_0(L; U)$ such that $x \geqq 0$. Then, from Proposition 2.1, $y = \eta x$ is a nontrivial nonnegative solution of

$$y(t) = -\int_a^b y(s) L_s^* H(t, s)\eta(t)/\eta(s)\, ds, \qquad y \in \mathscr{L}_p[a, b].$$

Therefore $y(t) \leqq \int_a^b y(s)[L_s^* H(t, s)\eta(t)/\eta(s)]_- \, ds$; taking the $\mathscr{L}_p$ norm of both sides of this expression and majorising the right-hand side by the Cauchy–Schwarz inequality gives a contradiction to (3.5).

COROLLARY 3.2.1. *Suppose* $G \in \mathscr{H}(U^*)$ *and* λ, ψ *are real-valued functions on* $[a, b]$ *with*

$$\psi \in \mathrm{AC}^{n-1}[a, b], \qquad U^*\psi = 0.$$

A sufficient condition for the conclusion of Theorem 3.2 *to hold is*

$$\sup \left\{ \int_a^b [L_s^* G(t, s) + \lambda(t) L^*\psi(s)]_- \, ds, \, a \leqq t \leqq b \right\} < 1. \tag{3.6}$$

This follows from the observation that if $H(t, s) = G(t, s) + \lambda(t)\psi(s)$, then $H \in \mathscr{H}(U^*)$. Corollary 3.2.1 relaxes the condition (3.1) of Theorem 3.1 since it is not necessary to assume $L^*\psi \geqq 0$ here. Indeed if $L^*\psi > 0$ then $\lambda(t)$ may be chosen so that $L^*H(t, \cdot) > 0$ except on a set of arbitrarily small measure. It can be seen from the following corollary that Theorem 3.2 is a generalization of Lyapunov's inequality [cf. 6, Theorem 5.1, p. 345].

COROLLARY 3.2.2. *Suppose there exists* $\psi \in \mathrm{AC}^1[a, b]$ *such that* $\psi(a)=\psi(b)=0$ *and*

$$\int_a^b \left[q_+(s)\frac{(b-s)(s-a)}{b-a} + \psi''(s) + q(s)\psi(s) \right]_+ ds \leqq 1. \tag{3.7}$$

Then there is no solution of

$$x'' + q(t)x = 0, \qquad x(a) = x(b) = 0$$

such that $x > 0$ *on* (a, b).

Proof. Consider the Green's function G defined by

$$-G(t, s)(b-a) = \begin{cases} (b-t)(s-a), & a \leqq s \leqq t, \\ (b-s)(t-a), & t \leqq s \leqq b. \end{cases}$$

Since $-(b-s)(s-a)/(b-a) < G(t, s) \leqq 0$ if $s \in (a, b)$, $s \neq t$, the condition (3.7) implies that (3.6) is satisfied with $\lambda(t) = -1$. The case $\psi = 0$ is Lyapunov's inequality. Observe however that, in contrast to the case $\psi = 0$, condition (3.7) need not in general imply the disconjugacy of $Ly = y'' + q(t)y$ on $[a, b]$ since the existence of ψ on $[a, b]$ does not necessarily imply the existence of a similar function for each subinterval of $[a, b]$.

THEOREM 3.3. *Suppose* $H \in \mathscr{H}(U^*)$ *and* $\eta \in \mathscr{N}_p(U)$ *exist such that*

$$\int_a^b \left\{ \int_a^b |L_s^* H(t, s)\eta(t)/\eta(s)|^q \, ds \right\}^{p/q} dt < 1, \tag{3.8}$$

$1 \leqq p \leqq \infty$, $1/p + 1/q = 1$. *Then Green's function exists for the problem* $\mathscr{P}(L; U)$. *If* p *or* $q = \infty$ *the corresponding integral in* (3.8) *should be replaced by the appropriate supremum.*

Proof. The condition (3.8) ensures that the map $\mathscr{T}: \mathscr{L}_p[a, b] \to \mathscr{L}_p[a, b]$ given by

$$(\mathscr{T}y)(t) = -\int_a^b y(s) L_s^* H(t, s)\eta(t)/\eta(s)\, ds$$

is a contraction and so $y = 0$ is the only fixed point. But if x is any solution of $\mathscr{P}_0(L; U)$, then $y = \eta x$ is a fixed point of $\mathscr{T}$ by Proposition 2.1. Thus $x = 0$ is the only solution of $\mathscr{P}_0(L; U)$, and Green's function for $\mathscr{P}(L; U)$ exists.

COROLLARY 3.3.1. *Suppose* $H \in \mathscr{H}(U^*)$ *is a Green's function which does not change sign on* $[a, b] \times [a, b]$ *and*

(i) $L^* H(t, \cdot) \leqq 0$ *for each* $t \in [a, b]$.

Suppose further that there exists $\phi \in \mathrm{AC}^{n-1}[a, b]$ *such that*

(ii) $U\phi = 0$, $1/|\phi| \in \mathscr{N}_\infty(U)$,

(iii) ϕ_0/ϕ *is continuous and positive on* $[a, b]$ *where*

$$\phi_0 = \int_a^b H(\cdot, s) L\phi(s)\, ds.$$

Then Green's function G *for* $\mathscr{P}(L; U)$ *exists,* G *and* H *have the same sign and*

$$|G(t, s)| \geqq |H(t, s)|.$$

Proof. First, Green's function G for $\mathscr{P}(L; U)$ exists since the formula (2.1) with $x = \phi$ implies

$$\phi(t) = \phi_0(t) - \int_a^b \phi(s) L_s^* H(t, s)\, ds,$$

and therefore from conditions (i), (ii)

$$0 \leqq -\frac{1}{\phi(t)}\int_a^b \phi(s) L_s^* H(t,s)\,ds \leqq 1-\frac{\phi_0(t)}{\phi(t)}.$$

Condition (iii) then gives

$$\sup\left\{\frac{1}{|\phi(t)|}\int_a^b |\phi(s)|\,|L_s^*H(t,s)|\,ds,\ a\leqq t\leqq b\right\}<1,$$

and the existence of Green's function for $\mathscr{P}(L;U)$ follows from Theorem 3.1 with $p=\infty$, $\eta=1/|\phi|$. Now the rest of the assertions follow by solving the equation

$$G(t,\tau)=H(t,\tau)-\int_a^b G(s,\tau)L_s^*H(t,s)\,ds \tag{3.9}$$

for G by Picard iterations. To verify the formula (3.9) consider $h=H(t,\cdot)$, $g=G(\cdot,\tau)$ so that

$$-\int_a^b gL^*h=\int_a^b (hLg-gL^*h)$$

$$=[hg](b)-[hg](\tau+)+[hg](\tau-)-[hg](t+)+[hg](t-)-[hg](a), \qquad \text{by (1.1)}$$

$$=-[hg](\tau+)+[hg](\tau-)-[hg](t+)+[hg](t-), \qquad \text{by (1.2),}$$

$$=-h(\tau)\{g^{(n-1)}(\tau+)-g^{(n-1)}(\tau-)\}+g(t)(-1)^n\{h^{(n-1)}(t+)-h^{(n-1)}(t-)\}$$

$$=-h(\tau)+g(t).$$

Remark. If instead of $H\in\mathscr{H}(U^*)$ it is assumed that $H(t,s)=K(s,t)$ where $(-1)^nK\in\mathscr{H}(U)$ and

(i)′ $LH(\cdot,s)\leqq 0$ for each $s\in[a,b]$, then it should be assumed that there exists a function $\psi\in \mathrm{AC}^{n-1}[a,b]$ such that

(ii)′ $U^*\psi=0$, $1/|\psi|\in\mathscr{N}_\infty(U^*)$,

(iii)′ ψ_0/ψ is continuous and positive on $[a,b]$ where

$$\psi_0=\int_a^b H(t,\cdot)L^*\psi(t)\,dt.$$

The conclusion of Corollary 3.3.1 then holds as before.

As an application for Corollary 3.3.1 consider Čaplygin's inequality which states that if $x\in\mathrm{AC}^{n-1}[a,b]$ and

$$x(a)=x'(a)=\cdots=x^{(n-1)}(a)=0, \qquad Lx\geqq 0,$$

then $x\geqq 0$ on an interval $[a,\gamma(a)]$. Clearly $[a,\gamma(a)]$ is the largest interval on which $K(s,t)\geqq 0$, $a\leqq s\leqq t\leqq\gamma(a)$ where K is the Gauchy function for L (i.e. the Green's function for the problem $\mathscr{P}(L;n,a,b)$). It is clear that $\gamma(a)\geqq\eta(a)$, where $\eta(a)$ is the first conjugate point of a.

COROLLARY 3.3.2. *Suppose H is a Cauchy function such that*

(i) $H(t,s)\geqq 0$, $a\leqq s\leqq t\leqq b$ *and*

(ii) *either* $L^*H(t,\cdot)\leqq 0$ *for each* t, $a\leqq t\leqq b$,

or $LH(\cdot,s)\leqq 0$ *for each* s, $a\leqq s\leqq b$.

Then the Cauchy function K for L satisfies

$$K(t,s)\geqq H(t,s)\geqq 0, \qquad a\leqq s\leqq t\leqq b.$$

Proof. It may be assumed that $a=0$, $b=1$. Let $(u_1, \cdots, u_n)$ be a fundamental solution set for $Lx=0$ such that $W(u_1, \cdots, u_n)>0$ on $[0, 1]$. Consider

$$\phi(t)=\mathscr{W}_{n+1}(u_1, \cdots, u_{n+1})\begin{bmatrix}0, & t\\ n, & 1\end{bmatrix}$$

where $u_{n+1}(t)=e^{\lambda t}$; it will be shown that if λ is large then

(3.10) $\phi(0)=\cdots=\phi^{(n-1)}(0)=0, \quad \phi^{(n)}(0)>0, \quad \phi>0 \quad \text{and} \quad L\phi>0 \quad \text{on } (0, 1]$,

and hence conditions (ii), (iii) of Corollary 3.3.1 are satisfied giving $K(t, s)\geqq H(t, s)\geqq 0$ if $L^*H(t, \cdot)\leqq 0$. To prove assertion (3.10), observe that

$$\begin{aligned}\phi(0)=\cdots&=\phi^{(n-1)}(0)=0,\\ \phi^{(n)}(0)&=W(u_1, \cdots, u_{n+1})(0)\\ &=W(u_1, \cdots, u_n)(0)\lambda^n+O(\lambda^{n-1}) \qquad (\lambda\to\infty).\end{aligned}$$

Also $L\phi=W(u_1, \cdots, u_n)(0)Lu_{n+1}>0$ for all large λ and thus $\phi>0$ on $(0, \eta(0)]$, where $\eta(0)$ is the first conjugate point of 0, by Čaplygin's inequality. In fact $\phi>0$ on $(0, 1]$ since, if $0<\eta(0)<1$, then

$$\phi(t)=W(u_1, \cdots, u_n)(0)\, e^{\lambda t}+O(\lambda^{n-1}) \qquad (\lambda\to\infty)$$

uniformly on $[\eta(0), 1]$. In the case that $LH(\cdot, s)\leqq 0$ an appropriate function ψ satisfying the conditions (ii)′, (iii)′ of the remark may be constructed in a similar fashion.

Corollary 3.3.2 may be combined with Pólya's Theorem V [13], cf. also [9], [10], to give a criterion other than disconjugacy for Čaplygin's inequality to hold. A particular case of this theorem states that if L_n is an operator which is disconjugate on $[c, d)$ then

$$f(c)=\cdots=f^{(k-1)}(c)=0, \qquad 1\leqq k\leqq n, \qquad L_nf\geqq 0 \quad \text{on } [c, d)$$

implies $L_{c,n-k}f\geqq 0$ on $[c, d]$ where $L_{c,0}y=y$ and, if $k<n$, $L_{c,n-k}$ is an operator of order $n-k$ the null set of which is $\{x: L_nx=0,\ x(c)=\cdots=x^{(k-1)}(c)=0\}$. It follows therefore that if H_c is the Cauchy function of L_n for initial value problems at c then, on $[c, d]$

$$L_{c,n-k}H_c(\cdot, s)\geqq 0, \qquad k=1, \cdots, n.$$

But if $c\leqq a$ then $H_a(\cdot, s)=H_c(\cdot, s)$ $a\leqq s\leqq t\leqq b$ so if $[a, b]\subset(c, d]$ and $q_k\in C^{n-k}(c, d]$, $q_k\leqq 0$, $h=1, \cdots, n$, then the function $H=H_a$ and the operator L, where

$$Ly=L_ny+\sum_{k=1}^{n} q_kL_{c,n-k}y,$$

satisfy the conditions of Corollary 3.3.2. The restriction $c<a$ is necessary since the operators $L_{c,n-k}$ are singular at c if $k<n$. The introduction to Levin's paper [8] contains a discussion of Čaplygin's inequality and disconjugacy and has further references to the literature on this subject.

COROLLARY 3.3.3. *Suppose that L is disconjugate on $[a, b]$ and that $\mathscr{P}(L; k, a, b)$ has Green's function H. If $q_n\in C[a, b]$, $(-1)^{n-k}q_n\leqq 0$. Then Green's function G for $\mathscr{P}(L+q_n; k, a, b)$ exists and*

$$(-1)^{n-k}G(t, s)\geqq(-1)^{n-k}H(t, s)\geqq 0, \qquad (t, s)\in[a, b]\times[a, b],$$

provided there exists a function $\phi \in AC^{n-1}[a, b]$ *having a zero of multiplicity* k *(exactly) at* a *and* $n-k$ *(exactly) at* b, $(-1)^{n-k}\phi>0$ *and* $(L+q_n)\phi \geqq 0$ *on* (a, b) *with strict inequality on a set of positive measure.*

Proof. Since ϕ has a zero of multiplicity k exactly at a (i.e. $\phi(a)=\cdots=\phi^{(k-1)}(a)=0$, $\phi^{(k)}(a)\neq 0$) and multiplicity $n-k$ exactly at b it follows that x/ϕ is bounded if $x(a)=\cdots=x^{(k-1)}(a)=x(b)=\cdots=x^{(n-k-1)}(b)=0$ and $x \in AC^{n-1}[a, b]$; thus $1/|\phi| \in \mathcal{N}_\infty(k; a, b)$. The Green's function H is continuous, $H(\cdot, s)$ has a zero of multiplicity k exactly at a and $n-k$ exactly at b and $(-1)^{n-k}H(\cdot, s)>0$ on (a, b), $a<s<b$ (cf. Coppel [3]). Therefore $\phi_0/\phi=1/\phi\int_a^b H(\cdot, s)(L+q_n)\phi(s)\,ds$ is continuous and positive on $[a, b]$. Also $(L+q_n)^*H(t, \cdot)=q_nH(t, \cdot)\leqq 0$ so Corollary 3.3.3 follows from Corollary 3.3.1.

Corollary 3.3.3 should be compared with Corollary 3.1.1 where it was assumed $(-1)^{n-k}q_n \geqq 0$ rather than $(-1)^{n-k}q_n \leqq 0$. Corollary 3.3.3 may be used to give another proof of the Levin–Nehari comparison theorem (Corollary 3.1.2) for operators $L+q_n$. This proof has the added advantage that it establishes the fact that the Green's function for $(L+q_n; k, a, b)$ depends monotonically on q_n.

By choosing H in Theorems 3.2, 3.3 to be specific Green's functions restrictions on $\|a_k\|_p$ may be found which guarantee the disconjugacy of L on $[a, b]$. However to obtain any significant improvement on results already in the literature it would appear necessary to obtain fairly good estimates on the quantities $(\partial^k/\partial s^k)H(t, s)$, $k=0, \cdots, n$. Estimates of this type on the Green's function for the problem $\mathcal{P}(D^n; k, a, b)$ may be found in the paper of Ostroumov [12]. It is hoped that this question may be considered further in a later paper.

The restriction $a_k \in C^{n-k}$ on the coefficients in L is not very important. The techniques adopted here may be extended to the generalized differential equations of Nehari [11] and the formal adjoint L^* of L may be discussed in this context. Many of the results may also be extended simply by approximation of coefficients which are continuous by functions which are smooth.

REFERENCES

[1] G. A. BESSMERTNYH AND A. JU. LEVIN, *Some inequalities satisfied by differentiable functions of one variable*, Dokl. Akad. Nauk SSSR, 144 (1962), pp. 471–474 = Soviet Math. Dokl., 3 (1962), pp. 737–740.

[2] EARL A. CODDINGTON AND NORMAN LEVINSON, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[3] W. A. COPPEL, *Disconjugacy*, Springer-Verlag, New York, 1971.

[4] PHILIP HARTMAN, *Principal solutions of disconjugate n-th order linear differential equations*, Amer. J. Math., 91 (1969), pp. 306–362; Corrigendum and addendum, Ibid., 93 (1971), pp. 439–451.

[5] ———, *Disconjugacy and Wronskians*, Japan–United States Seminar on Ordinary Differential and Functional Equations, M. Urabe, ed., Springer-Verlag, New York, 1971, pp. 208–218.

[6] ———, *Ordinary Differential Equations*, Hartman, Baltimore, 1973.

[7] A. JU. LEVIN, *Some problems bearing on the oscillation of solutions of linear differential equations*, Dokl. Akad. Nauk SSSR, 148 (1963), pp. 512–515 = Soviet Math. Dokl., 4 (1963), pp. 121–124.

[8] ———, *Non-oscillation of solutions of the equation* $x^{(n)}+p_1(t)x^{(n-1)}+\cdots+p_n(t)x=0$, Uspehi Mat. Nauk, 24 (1969), pp. 43–96 = Russian Math. Surveys, 24 (1969), pp. 43–99.

[9] J. S. MULDOWNEY, *On an inequality of Caplygin and Pólya*, Proc. Royal Irish Acad. Sect. A, 76 (1976), pp. 85–99.

[10] ———, *Linear differential inequalities*, this Journal, to appear.

[11] ZEEV NEHARI, *Disconjugate linear differential operators*, Trans. Amer. Math. Soc., 129 (1967), pp. 500–576.

[12] V. V. OSTROUMOV, *Unique solvability of the de la Vallée Poussin problem*, Differentsial'nye Uravneniya 4 (1968), pp. 261–268 = Differential Equations, 4 (1968), pp. 135–139.

[13] G. PÓLYA, *On the mean-value theorem corresponding to a given linear homogeneous differential equation*, Trans. Amer. Math. Soc., 24 (1922), pp. 312–324.
[14] WILLIAM T. REID, *Ordinary Differential Equations*, John Wiley, New York, 1971.
[15] THOMAS L. SHERMAN, *Properties of solutions of n-th order linear differential equations*, Pacific J. Math., 15 (1965), pp. 1045–1060.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0020 $01.00/0

ON THE OSCILLATORY AND ASYMPTOTIC BEHAVIOR OF A CLASS OF DIFFERENTIAL EQUATIONS WITH DEVIATING ARGUMENTS*

Y. G. SFICAS† AND I. P. STAVROULAKIS‡

Abstract. This paper is concerned with the oscillatory and asymptotic behavior of the solutions of the differential equation

$$(E, \delta) \qquad [r(t)x^{(n-1)}(t)]' + \delta \sum_{i=1}^{N} p_i(t) f_i(x[\tau_1(t)], \cdots, x[\tau_m(t)]) = 0$$

where $\delta = \pm 1$. Conditions which insure that every solution of (E, 1) is oscillatory or tending monotonically to zero are given. A classification of all solutions of (E, −1) with respect to their behavior as $t \to \infty$ and to their oscillatory character is also obtained. Finally a comparison with the oscillatory behavior of second order equations is presented. The obtained results unify, extend and improve recent ones by Lovelady and by the authors.

1. Introduction. This paper is concerned with the oscillatory and asymptotic behavior of the solutions of the differential equation with deviating arguments

$$(E, \delta) \qquad [r(t)x^{(n-1)}(t)]' + \delta \sum_{i=1}^{N} p_i(t) f_i(x[\tau_1(t)], \cdots, x[\tau_m(t)]) = 0, \qquad t \geqq t_0,$$

where $\delta = \pm 1$.

In the particular case of the second order linear differential equation

$$(1.1) \qquad x'' + p(t)x = 0,$$

Hille [1] (see also [11, p. 45]) obtained the following result:

If

$$g_* = \liminf_{t\to\infty} t \int_t^\infty p(s)\, ds \quad \text{and} \quad g^* = \limsup_{t\to\infty} t \int_t^\infty p(s)\, ds,$$

then the conditions $g_* \leqq 1/4$, $g^* \leqq 1$ *are necessary conditions and* $g^* < 1/4$ *is a sufficient condition for* (1.1) *to be nonoscillatory*; $g_* > 1/4$ *is a sufficient condition for* (1.1) *to be oscillatory.*

This result can be applied in many cases in which other results by Kiguradze [3] and by Sficas [7] fail.

For higher order differential equations with deviating arguments the first attempts known to the present authors in stating analogous results were made by Sficas [8] and Lovelady [4], [5], [6]. It must be noted that the results by Sficas can be applied only to retarded differential equations while those by Lovelady concern only linear differential equations (ordinary or retarded).

The purpose of the present paper is:

a) To extend the results by Lovelady in such a way that they can be applied in cases of nonlinear differential equations and especially with deviating arguments.

b) To include differential equations which contain a damping term.

Although the method for the general case of equations which contain more than one deviating argument is the same as for one argument, we preferred to deal with the more general case (though we are afraid that this will be tiring for the reader) because, as our results are given, it is clearer how each deviating argument acts upon the oscillatory behavior of equations.

* Received by the editors August 4, 1976 and in final revised form February 14, 1977.

† Department of Mathematics, University of Ioannina, Ioannina, Greece.

‡ Department of Mathematics, University of Ioannina, Ioannina, Greece. During the final preparation of this paper this author was supported by the National Hellenic Research Foundation.

2. Preliminaries. Consider the nth order $(n>1)$ differential equation (E, δ). We suppose that $\tau_j (j=1,2,\cdots,m)$, $p_i (i=1,2,\cdots,N)$, $f_i (i=1,2,\cdots,N)$ and r are continuous real-valued functions such that:

(i) τ_j *are defined on the half-line* $[t_0, \infty)$ *and*

$$\lim_{t\to\infty} \tau_j(t)=\infty \qquad (j=1,2,\cdots,m);$$

(ii) p_i *are nonnegative on* $[t_0,\infty)$;

(iii) f_i *are defined on* $\mathbf{R}^m$, *where* $\mathbf{R}$ *is the real line*

$$(\forall j=1,2,\cdots,m)\; y_j>0 \Rightarrow f_i(y_1,y_2,\cdots,y_m)>0$$

and

$$(\forall j=1,2,\cdots,m)\; y_j<0 \Rightarrow f_i(y_1,y_2,\cdots,y_m)<0;$$

(iv) r *is nonnegative on* $[t_0,\infty)$ *and such that*

$$\int^{\infty} \frac{dt}{r(t)}=\infty.$$

The above conditions will be assumed in the sequel without further mention.

Throughout this paper, by "solution" of the differential equation (E, δ) we shall mean only solutions x which are defined on the half-line $[t_x, \infty)$. The oscillatory character is considered in the usual sense, i.e. a solution is called *oscillatory* if it has no last zero, otherwise it is called *nonoscillatory*.

To obtain our results we need the following Lemma which is easily derived from two lemmas due to Kiguradze (cf. [2] and [9]).

LEMMA. *Suppose that* x *is a positive* $(n-1)$-*times continuously differentiable function on an interval* $[a,\infty)$ *such that the function* $r(t)x^{(n-1)}(t)$ *is continuously differentiable on* $[a,\infty)$, *where* r *satisfies* (iv). *Furthermore consider the functions* y_i *defined as follows*:

$$y_i=\begin{cases} x^{(i)}, & 0\leqq i\leqq n-2,\\ rx^{(n-1)}, & i=n-1,\\ [rx^{(n-1)}], & i=n.\end{cases}$$

If y_n *is of constant sign and not identically zero for all large* t, *then there exist a* $t_x\geqq a$ *and an integer* l, $0\leqq l\leqq n$ *with* $n+l$ *odd for* $y_n\leqq 0$, $n+l$ *even for* $y_n\geqq 0$, *and such that for every* $t\geqq t_x$

$$l>0 \Rightarrow y_i(t)>0 \qquad (i=0,1,\cdots,l-1)$$

and

$$l\leqq n-1 \Rightarrow (-1)^{l+i}\, y_i(t)>0 \qquad (i=l,l+1,\cdots,n-1).$$

The oscillatory and asymptotic behavior of bounded solutions of the differential equation (E, δ) is well described by the following theorem. The proof of this theorem is omitted since it can be proved as in the case when $r(t)\equiv 1$, with some appropriate modifications.

THEOREM 1. *Let the following condition be satisfied*:

(C_1) *For some* i, $1 \leqq i \leqq N$, *either*

$$\int^{\infty} p_i(t)\, dt = \infty$$

or

$$\int^{\infty} \frac{t^{n-2}}{r(t)} \int_t^{\infty} p_i(s)\, ds\, dt = \infty.$$

Then every bounded solution x *of* (E, 1) {*resp.* (E, -1)} is:

(a) *oscillatory for* n *even* {*resp. odd*},

(b) *for* n *odd* {*resp. even*}, *either oscillatory or tending monotonically to zero as* $t \to \infty$, *together with its first* $n-2$ *derivatives. In the latter case we also have* $\lim_{t\to\infty} r(t)x^{(n-1)}(t) = 0$.

3. Main results. In the following theorems we consider the functions f, τ_j^* and T_j defined as follows:

$$f(y_1, y_2, \cdots, y_m) = \min_{1 \leqq i \leqq N} |f_i(y_1, y_2, \cdots, y_m)|,$$

$$\tau_j^*(t) = \inf_{s \geqq t} (\min \{s, \tau_j(s)\}), \qquad T_j(t) = \inf_{s \geqq t} (\max \{s, \tau_j(s)\}) \qquad (j = 1, 2, \cdots, m).$$

We remark that these functions, as they have been defined, satisfy the following conditions:

(v) τ_j^* *are nondecreasing on* $[t_0, \infty)$,

$$\tau_j^*(t) \leqq t \quad \textit{for every } t \geqq t_0,$$

and

$$\lim_{t\to\infty} \tau_j^*(t) = \infty.$$

(vi) T_j *are nondecreasing on* $[t_0, \infty)$,

$$T_j(t) \geqq t \quad \textit{for every } t \geqq t_0,$$

and

$$\lim_{t\to\infty} T_j(t) = \infty.$$

Now we introduce the following conditions in which the functions τ and T are defined as follows:

$$\tau(t) = \min_{1 \leqq j \leqq m} \tau_j^*(t) \quad \text{and} \quad T(t) = \max_{1 \leqq j \leqq m} T_j(t).$$

(C_2) *There exist nonnegative numbers* α_j, $j = 1, 2, \cdots, m$ *with* $\sum_{j=1}^{m} \alpha_j = 1$ *and some*

$1 \leqq i \leqq N$ *such that for every integer* l, $1 \leqq l \leqq n-1$ *with* $n+l$ *odd*,

$$l<n-1 \Rightarrow \qquad \int^{\infty} p_i(s) \prod_{j=1}^{m} [\tau_j^*(s)]^{(l-1)\alpha_j}\, ds = \infty, \quad \text{or}$$

$$\int^{\infty} \frac{s^{n-l-2}}{r(s)} \int_s^{\infty} p_i(u) \prod_{j=1}^{m} [\tau_j^*(u)]^{(l-1)\alpha_j}\, du\, ds = \infty;$$

$$l=n-1 \Rightarrow \qquad \int^{\infty} p_i(u) \prod_{j=1}^{m} [\tau_j^*(u)]^{(n-2)\alpha_j}\, du = \infty.$$

(C_3) *There exist nonnegative numbers* α_j, $j=1,2,\cdots,m$ *with* $\sum_{j=1}^m \alpha_j = 1$ *such that for every integer* l, $1 \leqq l \leqq n-1$ *with* $n+l$ *odd*

$$l<n-1 \Rightarrow \limsup_{t\to\infty} \tau(t) \int_{T(t)}^{\infty} \frac{[s-T(t)]^{n-l-2}}{r(s)} \int_s^{\infty} \sum_{i=1}^{N} p_i(u) \prod_{j=1}^{m} [\tau_j^*(u)-\tau_j^*(t)]^{(l-1)\alpha_j}\, du\, ds > (l-1)!(n-l-2)!c;$$

$$l=n-1 \Rightarrow \limsup_{t\to\infty} \left[\int_{t_0}^{\tau(t)} \frac{ds}{r(s)} \right] \int_{T(t)}^{\infty} \sum_{i=1}^{N} p_i(u) \prod_{j=1}^{m} [\tau_j^*(u)-\tau_j^*(t)]^{(n-2)\alpha_j}\, du > (n-2)!c$$

where

$$c = \max\left\{ \limsup_{\substack{y_j\to\infty\\ 1\leqq j\leqq m}} \frac{y_1^{\alpha_1}\cdots y_m^{\alpha_m}}{f(y_1,\cdots,y_m)}, \limsup_{\substack{y_j\to-\infty\\ 1\leqq j\leqq m}} \frac{|y_1|^{\alpha_1}\cdots |y_m|^{\alpha_m}}{f(y_1,\cdots,y_m)} \right\}.$$

(C_2') *This condition is as* (C_2) *but the integer* l *is such that* $1 \leqq l \leqq n-2$ *with* $n+l$ *even.*

(C_3') *This condition is as* (C_3) *but the integer* l *is such that* $1 \leqq l \leqq n-2$ *with* $n+l$ *even.*

THEOREM 2. *Suppose that either condition* (C_2) *or condition* (C_3) *holds. Then every solution* x *of* (E, 1) *is*:

(a) *oscillatory, for* n *even,*

(b) *for* n *odd, either oscillatory or tending monotonically to zero as* $t\to\infty$, *together with its first* $n-2$ *derivatives. In the latter case we also have* $\lim_{t\to\infty} r(t)x^{(n-1)}(t)=0$.

Proof. Let x be a nonoscillatory solution of (E, 1) with $\lim_{t\to\infty} x(t) \neq 0$. Since the substitution $u=-x$ transforms (E, 1) into an equation of the same form satisfying the assumptions of the theorem we can suppose, without loss of generality, that $x(t)>0$ for every $t \geqq t_0$.

By (i), we can choose a $t_1 \geqq t_0$ such that for every $t \geqq t_1$

$$\tau_j(t) \geqq t_0 \qquad (j=1,2,\cdots,m).$$

Thus, by (E, 1), in view of (ii) and (iii), for every $t \geqq t_1$ we obtain

$$y_n(t) \equiv [r(t)x^{(n-1)}(t)]' = -\sum_{i=1}^{N} p_i(t) f_i(x[\tau_1(t)],\cdots,x[\tau_m(t)]) \leqq 0.$$

We notice that since the functions $p_i(t)$ ($i=1,2,\cdots,N$) are, by (C_2) and (C_3), not identically zero for all large t the same holds for $y_n(t)$. Thus, applying the Lemma, we derive that there exist a $t_x \geqq t_1$ and an integer l, $0 \leqq l \leqq n$ with $n+l$ odd such that for every $t \geqq t_x$

$$\text{(3.1)} \qquad \begin{aligned} &l>0 \Rightarrow y_i(t)>0 && (i=0,1,\cdots,l-1),\\ &l \leqq n-1 \Rightarrow (-1)^{l+i} y_i(t) > 0 && (i=l,l+1,\cdots,n-1). \end{aligned}$$

Since the integer $n+l$ is odd, obviously $l \leqq n-1$. Furthermore $l>0$. To prove this, we show first that conditions (C_2) and (C_3) imply (C_1). That condition (C_2) implies (C_1) is obvious. Now, by (C_3), we have

$$\begin{aligned} 0 &< \limsup_{t\to\infty} \tau(t) \int_{T(t)}^{\infty} \frac{[s-T(t)]^{n-3}}{r(s)} \int_s^{\infty} \sum_{i=1}^{N} p_i(u)\,du\,ds \\ &\leqq \limsup_{t\to\infty} t \int_{T(t)}^{\infty} \frac{s^{n-3}}{r(s)} \int_s^{\infty} \sum_{i=1}^{N} p_i(u)\,du\,ds \\ &\leqq \limsup_{t\to\infty} \int_t^{\infty} \frac{s^{n-2}}{r(s)} \int_s^{\infty} \sum_{i=1}^{N} p_i(u)\,du\,ds, \end{aligned}$$

and therefore there exists an i, $1 \leqq i \leqq N$ such that

$$\limsup_{t\to\infty} \int_t^{\infty} \frac{s^{n-2}}{r(s)} \int_s^{\infty} p_i(u)\,du\,ds > 0$$

which obviously implies (C_1). Thus, by Theorem 1, the solution x is not bounded and consequently, in view of the Lemma, $l>0$. Since $l>0$ and $l \leqq n-1$ for every $t \geqq t_x$ we have

$$\begin{aligned} & y_i(t) > 0 \qquad (i=0,1,\cdots,l-1), \\ & (-1)^{l+i} y_i(t) > 0 \qquad (i=l,l+1,\cdots,n-1). \end{aligned} \tag{3.2}$$

If $l<n-1$, then an integration of (E, 1) $n-l$ times from t to ∞ yields

$$\begin{aligned} (-1)^{n-l-1} x^{(l)}(t) \geqq \int_t^{\infty} \frac{(s-t)^{n-l-2}}{r(s)(n-l-2)!} \int_s^{\infty} \sum_{i=1}^{N} p_i(u) f_i(x[\tau_1(u)], \\ \cdots, x[\tau_m(u)])\,du\,ds, \qquad t \geqq t_x. \end{aligned} \tag{3.3}$$

Since the integer $n+l$ is odd, the above inequality for every $t \geqq t_x$ gives

$$\begin{aligned} x^{(l)}(t) &\geqq \int_{T(t)}^{\infty} \frac{[s-T(t)]^{n-l-2}}{r(s)(n-l-2)!} \int_s^{\infty} \sum_{i=1}^{N} p_i(u) f_i(x[\tau_1(u)],\cdots,x[\tau_m(u)])\,du\,ds \\ &= \int_{T(t)}^{\infty} \frac{[s-T(t)]^{n-l-2}}{r(s)(n-l-2)!} \int_s^{\infty} \sum_{i=1}^{N} p_i(u) \frac{f_i(x[\tau_1(u)],\cdots,x[\tau_m(u)])}{\prod_{i=1}^{m} x^{\alpha_i}[\tau_j(u)]} \sum_{j=1}^{m} x^{\alpha_i}[\tau_j(u)]\,du\,ds \\ &\geqq \inf_{u \geqq T(t)} \frac{f(x[\tau_1(u)],\cdots,x[\tau_m(u)])}{\prod_{j=1}^{m} x^{\alpha_i}[\tau_j(u)]} \int_{T(t)}^{\infty} \frac{[s-T(t)]^{n-l-2}}{r(s)(n-l-2)!} \\ &\quad \cdot \int_s^{\infty} \sum_{i=1}^{N} p_i(u) \prod_{j=1}^{m} x^{\alpha_i}[\tau_j(u)]\,du\,ds. \end{aligned}$$

Hence

$$\begin{aligned} x^{(l)}(t) \geqq \inf_{\substack{y_j \geqq x[\tau_j^*(T(t))] \\ 1 \leqq j \leqq m}} \frac{f(y_1,\cdots,y_m)}{\prod_{j=1}^{m} y_j^{\alpha_i}} \int_{T(t)}^{\infty} \frac{[s-T(t)]^{n-l-2}}{r(s)(n-l-2)!} \\ \cdot \int_s^{\infty} \sum_{i=1}^{N} p_i(u) \prod_{j=1}^{m} x^{\alpha_i}[\tau_j^*(u)]\,du\,ds. \end{aligned} \tag{3.4}$$

Consequently, by Taylor's formula with integral remainder

$$x(v) = \sum_{k=0}^{l-2} \frac{(v-w)^k x^{(k)}(w)}{k!} + \int_w^v \frac{(v-\xi)^{l-2} x^{(l-1)}(\xi)}{(l-2)!}\,d\xi, \qquad t_x \leqq w \leqq v. \tag{3.5}$$

Because $\lim_{t\to\infty} \tau_j^*(t)=\infty$, there exists a $t_2 \geqq t_x$ such that for every $t \geqq t_2$

$$\tau_j^*(t) \geqq t_x \qquad (j=1,2,\cdots,m)$$

and therefore (3.5), in view of (3.2), gives

$$x[\tau_j^*(u)] \geqq \int_{\tau_j^*(t)}^{\tau_j^*(u)} \frac{[\tau_j^*(u)-\xi]^{l-2}}{(l-2)!} x^{(l-1)}(\xi)\,d\xi, \qquad t_2 \leqq t \leqq u. \tag{3.6}$$

Combining (3.4) and (3.6), for every $t \geqq t_2$ we obtain

$$\begin{aligned} x^{(l)}(t) \geqq & \inf_{y_j \geqq x[\tau_j^*(T(t))]} \frac{f(y_1,\cdots,y_m)}{\prod_{j=1}^m y_j^{\alpha_j}} \int_{T(t)}^\infty \frac{[s-T(t)]^{n-l-2}}{r(s)(n-l-2)!} \int_s^\infty \sum_{i=1}^N p_i(u) \\ & \cdot \prod_{j=1}^m \left[\int_{\tau_j^*(t)}^{\tau_j^*(u)} \frac{[\tau_j^*(u)-\xi]^{l-2}}{(l-2)!} x^{(l-1)}(\xi)\,d\xi\right]^{\alpha_j} du\,ds \\ \geqq & \inf_{y_j \geqq x[\tau_j^*(T(t))]} \frac{f(y_1,\cdots,y_m)}{\prod_{j=1}^m y_j^{\alpha_j}} \int_{T(t)}^\infty \frac{[s-T(t)]^{n-l-2}}{r(s)(n-l-2)!} \int_s^\infty \sum_{i=1}^N p_i(u) \\ & \cdot \prod_{j=1}^m [x^{(l-1)}[\tau(t)]]^{\alpha_j} \left[\int_{\tau_j^*(t)}^{\tau_j^*(u)} \frac{[\tau_j^*(u)-\xi]^{l-2}}{(l-2)!}\,d\xi\right]^{\alpha_j} du\,ds. \end{aligned}$$

That is, for every $t \geqq t_2$

$$\begin{aligned} x^{(l)}(t) \geqq & \frac{x^{(l-1)}[\tau(t)]}{(l-1)!(n-l-2)!} \inf_{y_j \geqq x[\tau_j^*(T(t))]} \frac{f(y_1,\cdots,y_m)}{\prod_{j=1}^m y_j^{\alpha_j}} \\ & \int_{T(t)}^\infty \frac{[s-T(t)]^{n-l-2}}{r(s)} \int_s^\infty \sum_{i=1}^N p_i(u) \prod_{j=1}^m [\tau_j^*(u)-\tau_j^*(t)]^{(l-1)\alpha_j}\,du\,ds, \end{aligned} \tag{3.7}$$

so

$$\int_{T(t)}^\infty \frac{[s-T(t)]^{n-l-2}}{r(s)} \int_s^\infty \sum_{i=1}^N p_i(u) \prod_{j=1}^m [\tau_j^*(u)-\tau_j^*(t)]^{(l-1)\alpha_j}\,du\,ds < \infty,$$

which implies the failure of (C_2).

If $l=n-1$ we have

$$x^{(n-1)}(t) \geqq \frac{1}{r(t)} \int_t^\infty \sum_{i=1}^N p_i(u) f_i(x[\tau_1(u)],\cdots,x[\tau_m(u)])\,du,$$

and following the preceding procedure we obtain

$$\begin{aligned} x^{(n-1)}(t) \geqq & \frac{x^{(n-2)}[\tau(t)]}{(n-2)!} \inf_{y_j \geqq x[\tau_j^*(T(t))]} \frac{f(y_1,\cdots,y_m)}{\prod_{j=1}^m y_j^{\alpha_j}} \\ & \cdot \frac{1}{r(t)} \int_{T(t)}^\infty \sum_{i=1}^N p_i(u) \prod_{j=1}^m [\tau_j^*(u)-\tau_j^*(t)]^{(n-2)\alpha_j}\,du, \qquad t \geqq t_2, \end{aligned} \tag{3.8}$$

thus

$$\int_{T(t)}^\infty \sum_{i=1}^N p_i(u) \prod_{j=1}^m [\tau_j^*(u)-\tau_j^*(t)]^{(n-2)\alpha_j}\,du < \infty,$$

which implies the failure of (C_2).

Suppose now that (C_2) fails. If $l<n-1$, *then* (3.7), *because* $x^{(l)}[\tau(t)] \geqq x^{(l)}(t)$, yields

$$\tau(t)\int_{T(t)}^{\infty}\frac{[s-T(t)]^{n-l-2}}{r(s)}\int_{s}^{\infty}\sum_{i=1}^{N}p_i(u)\prod_{j=1}^{m}[\tau_j^*(u)-\tau_j^*(t)]^{(l-1)\alpha_j}\,du\,ds$$
(3.9)
$$\leqq (l-1)!(n-l-2)!\sup_{y_j\geqq x[\tau_j^*(T(t))]}\frac{y_1^{\alpha_1}\cdots y_m^{\alpha_m}}{f(y_1,\cdots,y_m)}\tau(t)\frac{x^{(l)}[\tau(t)]}{x^{(l-1)}[\tau(t)]}, \qquad t\geqq t_2.$$

It is a matter of elementary calculus to show that

$$\limsup_{t\to\infty}\tau(t)\frac{x^{(l)}[\tau(t)]}{x^{(l-1)}[\tau(t)]}\leqq 1. \tag{3.10}$$

On the other hand $\lim_{t\to\infty} x(t)=\infty$ and therefore

$$\lim_{t\to\infty}\left\{\sup_{y_j\geqq x[\tau_j^*(T(t))]}\frac{y_1^{\alpha_1}\cdots y_m^{\alpha_m}}{f(y_1,\cdots,y_m)}\right\}=\limsup_{t\to\infty}\frac{y_1^{\alpha_1}\cdots y_m^{\alpha_m}}{f(y_1,\cdots,y_m)}\leqq c.$$

Combining (3.9), (3.10) and the last inequality we obtain

$$\limsup_{t\to\infty}\tau(t)\int_{T(t)}^{\infty}\frac{[s-T(t)]^{n-l-2}}{r(s)}\int_{s}^{\infty}\sum_{i=1}^{N}p_i(u)\prod_{j=1}^{m}[\tau_j^*(u)-\tau_j^*(t)]^{(l-1)\alpha_j}\,du\,ds \leqq (l-1)!(n-l-2)!c,$$

which contradicts (C_3).

If $l=n-1$ then, by (3.8), since $r[\tau(t)]x^{(n-1)}[\tau(t)]\geqq r(t)x^{(n-1)}(t)$, we have

$$\int_{T(t)}^{\infty}\sum_{i=1}^{N}p_i(u)\prod_{j=1}^{m}[\tau_j^*(u)-\tau_j^*(t)]^{(n-2)\alpha_j}\,du \leqq (n-2)!\sup_{y_j\geqq x[\tau_j^*(T(t))]}\frac{y_1^{\alpha_1}\cdots y_m^{\alpha_m}}{f(y_1,\cdots,y_m)}\frac{r[\tau(t)]x^{(n-1)}[\tau(t)]}{x^{(n-2)}[\tau(t)]}.$$

Furthermore, because the function $r(t)x^{(n-1)}(t)$ is nonincreasing, for every $t\geqq t_2$ we get

$$\frac{x^{(n-2)}[\tau(t)]}{r[\tau(t)]x^{(n-1)}[\tau(t)]}=\frac{x^{(n-2)}[\tau(t_2)]+\int_{\tau(t_2)}^{\tau(t)}(1/r(s))[r(s)x^{(n-1)}(s)]\,ds}{r[\tau(t)]x^{(n-1)}[\tau(t)]}$$
$$\geqq\frac{x^{(n-2)}[\tau(t_2)]+r[\tau(t)]x^{(n-1)}[\tau(t)]\int_{\tau(t)_2}^{\tau(t)}(ds/r(s))}{r[\tau(t)]x^{(n-1)}[\tau(t)]}\geqq\int_{\tau(t_2)}^{\tau(t)}\frac{ds}{r(s)}.$$

Hence

$$\limsup_{t\to\infty}\left[\int_{\tau(t_2)}^{\tau(t)}\frac{ds}{r(s)}\right]\int_{T(t)}^{\infty}\sum_{i=1}^{N}p_i(u)\prod_{j=1}^{m}[\tau_j^*(u)-\tau_j^*(t)]^{(n-2)\alpha_j}\,du\leqq(n-2)!c,$$

which contradicts (C_3).

We therefore conclude that every solution x of (E, 1) is either oscillatory or such that $\lim_{t\to\infty} x(t)=0$. In the last case, by the mean value theorem, we get

$$\lim_{t\to\infty} x^{(k)}(t)=0 \qquad (k=0,1,\cdots,n-2)$$

and, by the Lemma, $l=0$ and therefore n is odd. Moreover, it is easily verified that $\lim_{t\to\infty} r(t)x^{(n-1)}(t)=0$.

THEOREM 3. *Let $\tau(t)$ be continuously differentiable. Suppose that there exist nonnegative numbers α_j, $j=1,2,\cdots,m$ with $\sum_{j=1}^m \alpha_j=1$ such that*

$$\mu=\frac{1}{2}\min\left\{2, \liminf_{y_j\to\infty}\frac{f(y_1,\cdots,y_m)}{\prod_{j=1}^m y_j^{\alpha_j}}, \liminf_{y_j\to-\infty}\frac{f(y_1,\cdots,y_m)}{\prod_{j=1}^m |y_j|^{\alpha_j}}\right\}$$

is a positive real number. If condition (C₂) *fails but either* a): *every solution of*

$$[r(t)w'(t)]'+\left\{\frac{\mu\tau'(t)}{(n-3)!}\int_{T(t)}^{\infty}[\tau(u)-\tau(t)]^{n-3}\sum_{i=1}^N p_i(u)\,du\right\}w[\tau(t)]=0 \tag{3.11}$$

is oscillatory, or b): *for any l, $0<l<n-1$ with $n+l$ odd, every solution of*

$$\begin{aligned} w''(t)+\Bigg\{&\frac{\mu}{(l-1)!(n-l-3)!}\int_{T(t)}^{\infty}\frac{[s-T(t)]^{n-l-3}}{r(s)}\\ &\cdot\int_s^{\infty}\sum_{i=1}^N p_i(u)\prod_{j=1}^m[\tau_j^*(u)-\tau_j^*(t)]^{(l-1)\alpha_j}\,du\,ds\Bigg\}w[\tau(t)]=0 \end{aligned} \tag{3.12}$$

is oscillatory, then the conclusion of Theorem 2 *holds.*

Proof. Consider again a nonoscillatory solution x of (E, 1) with $\lim_{t\to\infty} x(t)\neq 0$. If $l=n-1$

$$\begin{aligned} r(t)x^{(n-1)}(t) &\geqq \inf_{y_j\geqq x[\tau_j^*(T(t))]}\frac{f(y_1,\cdots,y_m)}{\prod_{j=1}^m y_j^{\alpha_j}}\int_{T(t)}^{\infty}\sum_{i=1}^N p_i(u)\prod_{j=1}^m x^{\alpha_j}[\tau_j(u)]\,du\\ &\geqq \mu\int_{T(t)}^{\infty}\sum_{i=1}^N p_i(u)x[\tau(u)]\,du\\ &\geqq \mu\int_{T(t)}^{\infty}\sum_{i=1}^N p_i(u)\int_{\tau(t)}^{\tau(s)}\frac{[\tau(u)-\xi]^{n-3}}{(n-3)!}x^{(n-2)}(\xi)\,d\xi\,du. \end{aligned}$$

Thus [6, Thm. 7] there is a continuous real-valued function w such that

$$r(t)w'(t)=\frac{\mu}{(n-3)!}\int_{T(t)}^{\infty}\sum_{i=1}^N p_i(u)\int_{\tau(t)}^{\tau(s)}[\tau(u)-\xi]^{n-3}w(\xi)\,d\xi\,du,\qquad t\geqq t_2. \tag{3.13}$$

Now differentiation of (3.13) yields (3.11), so w is a nonoscillatory solution of (3.11). This completes the proof if $l=n-1$.

If $l<n-1$, then following a procedure similar to that of the proof of Theorem 2 we obtain

$$\begin{aligned} -x^{(l+1)}(t)\geqq \mu&\frac{x^{(l-1)}[\tau(t)]}{(l-1)!(n-l-3)!}\int_{T(t)}^{\infty}\frac{[s-T(t)]^{n-l-3}}{r(s)}\\ &\cdot\int_s^{\infty}\sum_{i=1}^N p_i(u)\prod_{j=1}^m[\tau_j^*(u)-\tau_j^*(t)]^{(l-1)\alpha_j}\,du\,ds,\qquad t\geqq t_2. \end{aligned}$$

If $\varphi(t)$ is given by

$$\varphi(t)=\frac{\mu}{(l-1)!(n-l-3)!}\int_{T(t)}^{\infty}\frac{[s-T(t)]^{n-l-3}}{r(s)}\int_s^{\infty}\sum_{i=1}^N p_i(u)\prod_{j=1}^m[\tau_j^*(u)-\tau_j^*(t)]^{(l-1)\alpha_j}\,du\,ds,$$

we get

$$x^{(l+1)}(t)+\varphi(t)x^{(l-1)}[\tau(t)]\leqq 0.$$

Thus we have a positive solution of

$$v''(t)+\varphi(t)v[\tau(t)]\leqq 0, \qquad t\geqq t_2.$$

Now the remainder of the proof proceeds as in [6, Thm. 7].

THEOREM 4. *Suppose that*

(C_4) *there exists some* i, $1\leqq i\leqq N$ *such that the function* $f_i(y_1,\cdots,y_m)$ *is nondecreasing in every variable* $y_1,\cdots,y_m$ *and for every* $M\neq 0$

$$\int^{\infty} p_i(s)f_i(M[\tau_1(s)]^{n-1},\cdots,M[\tau_m(s)]^{n-1})\,ds=\pm\infty,$$

and either condition (C_2') *or* (C_3') *holds. Then every solution* x *of* (E, -1) *satisfies exactly one of the following*:

(I) x *is oscillatory*,

(II) x *and its first* $n-2$ *derivatives tend monotonically to zero as* $t\to\infty$ *and moreover* $\lim_{t\to\infty} r(t)x^{(n-1)}(t)=0$;

(III) *either*

$$\lim_{t\to\infty} r(t)x^{(n-1)}(t)=\infty, \qquad \lim_{t\to\infty} x^{(j)}(t)=\infty \qquad (j=0,1,\cdots,n-2)$$

or

$$\lim_{t\to\infty} r(t)x^{(n-1)}(t)=-\infty, \qquad \lim_{t\to\infty} x^{(j)}(t)=-\infty \qquad (j=0,1,\cdots,n-2).$$

Moreover (II) *occurs only in the case of even* n.

Proof. Let x be a nonoscillatory solution of (E, -1) with $\lim_{t\to\infty} x(t)\neq 0$. As in the proof of Theorem 2, we suppose, without loss of generality, that $x(t)>0$ for every $t\geqq t_0$. Moreover, we can choose $t_1\geqq t_0$ such that

$$y_n(t)\equiv[r(t)x^{(n-1)}(t)]'\geqq 0$$

for every $t\geqq t_1$, where $y_n(t)$ is not identically zero for all large t. Thus, by the Lemma, there exists $t_x\geqq t_1$ and an integer l, $0\leqq l\leqq n$ with $n+l$ even, such that for every $t\geqq t_x$ the relation (3.1) holds.

Consequently we consider the following two cases:

Case 1: $l=n$. In this case we have

$$y_i(t)>0 \quad \text{for every } t\geqq t_x \qquad (i=0,1,\cdots,n-1). \tag{3.14}$$

Furthermore, by Taylor's formula

$$x(t)\geqq x(t_x)+\frac{(t-t_x)x'(t_x)}{1!}+\cdots+\frac{(t-t_x)^{n-1}x^{(n-1)}(t_x)}{(n-1)!},$$

and therefore there exists a constant $M>0$ and $t_2\geqq t_x$ such that for every $t\geqq t_2$,

$$x(t)\geqq Mt^{n-1}. \tag{3.15}$$

Integrating equation (E, 1) from t_3 to t we obtain

$$r(t)x^{(n-1)}(t)=r(t_3)x^{(n-1)}(t_3)+\int_{t_3}^{t}\sum_{i=1}^{N} p_i(s)f_i(x[\tau_1(s)],\cdots,x[\tau_m(s)])\,ds$$

where $t_3\geqq t_2$ has been chosen so that for every $t\geqq t_3$

$$\tau_j(t)\geqq t_2 \qquad (j=1,2,\cdots,m).$$

Thus, by (3.14), (C_4) and (3.15), we get

$$r(t)x^{(n-1)}(t) \geqq \int_{t_3}^{t} \sum_{i=1}^{N} p_i(s) f_i(M[\tau_1(s)]^{n-1}, \cdots, M[\tau_m(s)]^{n-1})\, ds,$$

and therefore

$$\lim_{t\to\infty} r(t)x^{(n-1)}(t) = \infty.$$

From this, it is easily shown that

$$\lim_{t\to\infty} x^{(j)}(t) = \infty \qquad (j=0, 1, \cdots, n-2).$$

Hence the solution x satisfies (III).

Case 2: $l < n-1$. In this case

$$y_{n-1}(t) < 0 \quad \text{for every } t \geqq t_x.$$

As in the proof of Theorem 2, it can be shown that $l > 0$. Hence the relations (3.2) and (3.3) hold. Then, following step-by-step the proof of Theorem 2 we again obtain a contradiction.

Hence, every nonoscillatory solution x of (E, -1) satisfies (III) or $\lim_{t\to\infty} x(t) = 0$. In the last case, by the mean value theorem, we get

$$\lim_{t\to\infty} x^{(k)}(t) = 0 \qquad (k=0, 1, \cdots, n-2)$$

and moreover $\lim_{t\to\infty} r(t)x^{(n-1)}(t) = 0$, that is, x satisfies (II). Hence, every solution x of (E, -1) satisfies exactly one of the properties (I), (II), (III). Finally, by the Lemma, we conclude that (II) occurs only in the case of even n.

THEOREM 5. *Suppose that μ is as in Theorem 3. If condition* (C_2') *fails but either* a): *every solution of*

$$w''(t) + \left\{\frac{\mu}{(n-4)!} \int_{T(t)}^{\infty} \frac{1}{r(s)} \int_s^{\infty} \sum_{i=1}^{N} p_i(u) \prod_{j=1}^{m} [\tau_j^*(u) - \tau_j^*(t)]^{(n-4)\alpha_j}\, du\, ds\right\} w[\tau(t)] = 0$$

is oscillatory, *or* b): *for any l, $0 < l < n-2$ with $n+l$ even, every solution of* (3.12) *is oscillatory, then every solution x of* (E, -1) *which does not satisfy* (III) *of Theorem 4 is*:

(a) *oscillatory for n odd,*

(b) *for n even, either oscillatory or tending monotonically to zero as $t \to \infty$, together with its first $n-2$ derivatives, and* $\lim_{t\to\infty} r(t)x^{(n-1)}(t) = 0$.

Proof. Let x be a nonoscillatory solution of (E, -1) with $\lim_{t\to\infty} x(t) \neq 0$, which does not satisfy (III) of Theorem 4. Here $n+l$ is even and $l < n$, hence $l \leqq n-2$.

If $l = n-2$

$$\begin{aligned} -x^{(n-2)}(t) &\geqq \inf_{y_j \geqq x[\tau_j^*(T(t))]} \frac{f(y_1, \cdots, y_m)}{\prod_{j=1}^{m} y_j^{\alpha_j}} \int_{T(t)}^{\infty} \frac{1}{r(v)} \int_v^{\infty} \sum_{i=1}^{N} p_i(u) \prod_{j=1}^{m} x^{\alpha_j}[\tau_j(u)]\, du\, dv \\ &\geqq \frac{\mu x^{(n-4)}[\tau(t)]}{(n-4)!} \int_{T(t)}^{\infty} \frac{1}{r(v)} \int_v^{\infty} \sum_{i=1}^{N} p_i(u) \prod_{j=1}^{m} [\tau_j^*(u) - \tau_j^*(t)]^{(n-4)\alpha_j}\, du\, dv. \end{aligned}$$

If $\varphi(t)$ is given by

$$\varphi(t) = \frac{\mu}{(n-4)!} \int_{T(t)}^{\infty} \frac{1}{r(v)} \int_v^{\infty} \sum_{i=1}^{N} p_i(u) \prod_{j=1}^{m} [\tau_j^*(u) - \tau_j^*(t)]^{(n-4)\alpha_j}\, du\, dv,$$

we get

$$x^{(n-2)}(t) + \varphi(t) x^{(n-4)}[\tau(t)] \leqq 0.$$

Thus we have a positive solution of

$$v''(t)+\varphi(t)v[\tau(t)]\leqq 0.$$

Now the proof proceeds as in [6, Thm. 7].

If $l<n-2$ then the proof is so similar to the last part of the proof of Theorem 3 that we omit it.

Remark. One can draw corollaries from Theorems 2–5 concerning the retarded differential equation

$$x^{(n)}(t)+\delta p(t)f(x[g(t)])=0, \qquad t\geqq t_0,$$

which extend recent results due to Lovelady [4], [5], [6]. Also, if $r(t)\equiv 1$ Theorems 2 and 4 yield recent results due to Stavroulakis [10].

REFERENCES

[1] E. HILLE, *Non-oscillation theorems*, Trans. Amer. Math. Soc., 64 (1948), pp. 234–252.

[2] I. T. KIGURADZE, *The problem of oscillation of solutions of nonlinear differential equations*, Differential'nye Uravnenija, 1 (1965), pp. 995–1006.

[3] ———, *Oscillation properties of solutions of certain ordinary differential equations*, Soviet Math., 3 (1962), pp. 649–652.

[4] D. L. LOVELADY, *Oscillation and even order linear differential equations*, Rocky Mountain J. Math., 6 (1976), pp. 299–304.

[5] ———, *Oscillation and a class of odd order linear differential equations*, Hiroshima Math. J., 5 (1975), pp. 371–383.

[6] ———, *Oscillation and a class of linear delay differential equations*, Trans. Amer. Math. Soc., 226 (1977), pp. 345–364.

[7] Y. G. SFICAS, *On oscillation and asymptotic behavior of a certain class of differential equations with retarded argument*, Utilitas Math., 3 (1973), pp. 239–249.

[8] ———, *Retarded actions on oscillations*, Bull. Soc. Math. Grèce, 17 (1976), pp. 1–10.

[9] ———, *On the oscillatory and asymptotic behavior of damped differential equations with retarded argument*, Hiroshima Math. J., 6 (1976), pp. 429–450.

[10] I. P. STAVROULAKIS, *Differential equations with deviating arguments—Contribution to the study of oscillatory and asymptotic properties of the solutions*, in Greek, Ph.D. thesis, University of Ioannina, 1976.

[11] C. A. SWANSON, *Comparison and Oscillation Theory of Linear Differential Equations*, Academic Press, New York, 1968.

SIAM J. MATH. ANAL.
Vol. 9, No. 5, October 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0905-0021 $01.00/0

ON A NONSELF ADJOINT EIGENFUNCTION EXPANSION*

D. NAYLOR†

Abstract. This paper constructs a formula of inversion of an integral transform similar to that associated with the names of Kontorovich and Lebedev but involving a truncated infinite interval. The transform in question is useful in the solution of certain exterior boundary value problems involving a complex boundary condition and the damped wave equation and is of some independent interest in that, despite the singular nonself adjoint nature of the underlying expansion problem and the complex nature of the eigenvalues, the formula of inversion can be expressed as an eigenfunction expansion which does not require a summability factor but which is convergent in the ordinary sense.

1. Introduction. In this paper a formula of inversion is constructed for the integral transform defined by the equation

$$F(u)=\int_a^\infty f(r)K_u(kr)\frac{dr}{r}. \tag{1}$$

Here k, a are positive constants and $K_u(kr)$ denotes the MacDonald type Bessel function, the notation being that of Watson [10]. The transform in question arises from a consideration of the differential equation

$$r^2y_{rr}+ry_r-(k^2r^2+u^2)y=0, \qquad a\leqq r<\infty, \tag{2}$$

together with the boundary condition

$$hy(a)+ay'(a)=0, \tag{3}$$

where $h=h_1+ih_2$ denotes a complex constant. The eigenfunctions are the Bessel functions $K_{u_n}(kr)$ where $u_1, u_2, \cdots$ are the zeros of the function $g(u)$ defined by the equation

$$g(u)=hK_u(ka)+kaK'_u(ka) \tag{4}$$

regarded as a function of u.

If the constant h is real the expansion problem though singular is self adjoint and the possible values of u^2 are real. It is shown in [3] that when h is real the function $g(u)$ possesses an infinite number of pairs of imaginary zeros $\pm u_n$ and at most one pair $\pm u_0$ of real zeros. In the self adjoint case the corresponding theory of such expansions as outlined in [9] justifies the existence of the expansion

$$f(r)=-2\sum_{u=u_n}\frac{u[hI_u(ka)+kaI'_u(ka)]K_u(kr)F(u)}{g'(u)} \tag{5}$$

where the summation includes all the zeros u_n located on the positive imaginary axis and, if present, the real positive zero u_0.

In this paper it will be proved that a formula of the type (5) is still valid even when the constant h is complex. Although the boundary problem (2), (3) that generates the expansion is, for complex h, both singular and nonself adjoint, the resulting series requires no summability factor to render it convergent. This situation is in contrast to that which prevails in the related singular and nonself adjoint expansion problem in which the kernel of the transform (1) is taken to be the Hankel function $H_u^{(1)}(kr)$

* Received by the editors October 14, 1976, and in final revised form May 18, 1977.

† Department of Applied Mathematics, University of Western Ontario, London, Ontario, Canada N6A 5B9.

rather than the Bessel function $K_u(kr)$, where k is real. This alternative choice leads to the transform

$$G(u)=\int_a^\infty H_u^{(1)}(kr)f(r)\frac{dr}{r}.$$

The formula of inversion of this transform cannot in general be expressed as a simple convergent series as in (5) but must be represented either as a contour integral or a series both of which involve a summability factor of one kind or another to render them convergent. The Hankel function series expansion was discussed in [1], [8] where it is pointed out that the resulting series is usually divergent and the problem of inverting the transform $G(u)$ was considered in detail in the author's papers [4], [5], [6], [7] where various alternative formulas of inversion are developed.

To discuss the validity or otherwise of an expansion like (5) in the nonself adjoint case it is necessary first of all to investigate the zeros of the function $g(u)$ when the constant h is complex. We shall prove that, for $h_2\neq 0$, the zeros are neither real nor purely imaginary. If $h_2>0$ the zeros are located in the first and third quadrants of the complex u-plane whilst if $h_2<0$ they lie in the second and fourth quadrants. An asymptotic formula is also obtained from which it is shown that the zeros approach the imaginary axis as $u_n\to\infty$. Figure 1 illustrates the disposition of the zeros in the case $h_2>0$.

2. The distribution of the eigenvalues. The following classical type argument, similar to that used in [2], can be followed to investigate the zeros in question. We write $y=K_u(kr)$ and multiply the equation (2) by $r^{-1}\bar{y}$ where $\bar{y}$ denotes the complex conjugate. Upon integrating the resulting equation we find that

$$u^2\int_a^\infty |y|^2\frac{dr}{r}=-\int_a^\infty [k^2|y|^2+|y_r|^2]r\,dr-a\bar{y}(a)y'(a), \tag{6}$$

the last term being obtained after an integration by parts. If we set $u=s+it$, $h=h_1+ih_2$ and separate the real and imaginary parts we find on using the condition (3) that

$$(s^2-t^2)\int_a^\infty |y|^2\frac{dr}{r}=-\int_a^\infty [k^2|y|^2+|y_r|^2]r\,dr+h_1|y(a)|^2, \tag{7}$$

$$2st\int_a^\infty |y|^2\frac{dr}{r}=h_2|y(a)|^2. \tag{8}$$

It follows immediately from (8) that if $h_2\neq 0$ there are no real zeros and no imaginary zeros. If $h_2>0$ the product st must also be positive which implies that the zeros are located in the first and third quadrants of the complex u-plane. However if $h_2<0$ the zeros must lie in the second and fourth quadrants.

Further information can be deduced as follows. The equation (7) implies the inequality

$$\begin{aligned}(s^2-t^2)\int_a^\infty |y|^2\frac{dr}{r}&\leqq -k^2\int_a^\infty |y|^2r\,dr+h_1|y(a)|^2\\ &\leqq -k^2a^2\int_a^\infty |y|^2\frac{dr}{r}+h_1|y(a)|^2\end{aligned} \tag{9}$$

since $r \geqq a^2/r$. If the integral surviving in (8), (9) is now eliminated we find that

$$s^2 - t^2 + k^2a^2 \leqq 2st \sinh \lambda,$$

that is

$$(t + s\,e^{\lambda})(t - s\,e^{-\lambda}) \geqq k^2a^2 \tag{10}$$

where we have set $h_1 = h_2 \sinh \lambda$, where λ is real. If equality is taken in (10) the resulting equation represents a hyperbola in the (s, t) plane whose asymptotes are the lines $s - t\,e^{\lambda} = 0$ and $s + t\,e^{-\lambda} = 0$. The two branches C_1 and C_2 of this curve are illustrated in Fig. 1 and the inequality (10) implies that the zeros of the function $g(u)$ must be located above C_1 and below C_2. If $h_2 > 0$ the zeros are of necessity situated in the first and third quadrants and so in this case they must lie in the regions shown in Fig. 1. Since $K_u(kr)$ is an even function of u then to every zero u_n there corresponds another zero $-u_n$, which gives rise to the same eigenfunction. This is illustrated in the figure. The hyperbolas intersect the imaginary axis where $t = \pm ka$ so it is noted that, when $h_2 > 0$, all the zeros are such that $|\mathrm{Im}\,(u)| \geqq ka$.

The asymptotic distribution of the zeros for large u can be obtained with the aid of the formulas

$$K_u(x) = \frac{\pi}{2 \sin u\pi}[I_{-u}(x) - I_u(x)], \tag{11}$$

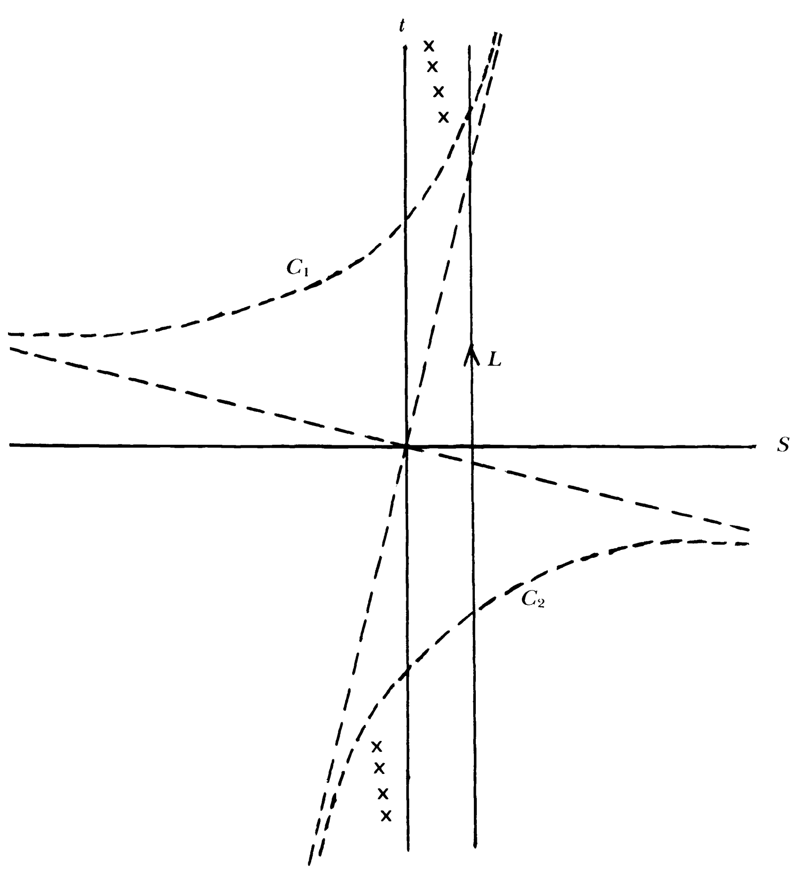

FIG. 1.

where

$$I_u(x)=\frac{(x/2)^u}{\Gamma(1+u)}\left[1+\frac{x^2}{4(1+u)}+O(u^{-2})\right]. \tag{12}$$

Formula (12) is valid for fixed x and large u bounded away from the negative integers. Upon inserting (11), (12) into (4) it can be shown after some reduction that the large zeros of (4) satisfy the equation

$$(ka/2)^{2u}=-\frac{\Gamma(1+u)}{\Gamma(1-u)}\left[1-\frac{k^2a^2+4h}{2u}+O(u^{-2})\right]. \tag{13}$$

The asymptotic behavior of the Γ-functions appearing in the preceding equation may be estimated with the aid of Stirling's formula

$$\Gamma(u)=(2\pi/u)^{1/2}\,e^{u\log u-u}\left[1+\frac{1}{12u}+O(u^{-2})\right] \tag{14}$$

which applies as $u\to\infty$ in the domain $|\arg u|<\pi$.

For definiteness we shall now assume that $h_2>0$. The relevant zeros now occur in the first quadrant and for them $\operatorname{Im}(u)\to+\infty$ so that $\sin u\pi\sim(i/2)\,e^{-iu\pi}$ as $u\to\infty$. It then follows from (14) in conjunction with the identity $\Gamma(u)\Gamma(1-u)=\pi\operatorname{cosec}u\pi$ that

$$\log\frac{\Gamma(1+u)}{\Gamma(1-u)}=2u\log(u/e)-iu\pi+\frac{i\pi}{2}+\frac{1}{6u}+O(u^{-2}) \tag{15}$$

as $u\to\infty$ in the first quadrant. Upon taking the logarithm of each side of (13) and inserting the expression (15) it follows that

$$u\left[\log(2u/(kae))-\frac{i\pi}{2}\right]=\frac{Q}{u}+\left(n+\frac{1}{4}\right)i\pi+O(u^{-2}) \tag{16}$$

where $Q=(3k^2a^2+12h-1)/12$ and n is a large positive integer.

To solve the equation (16) we set $u=\mathrm{Re}^{i\theta}$ where $0\leqq\theta\leqq\pi/2$ and equate real and imaginary parts. This leads to the equations

$$R\cos\theta\log(2R/(kae))+R(\pi/2-\theta)\sin\theta=\frac{Q_1\cos\theta+h_2\sin\theta}{R}+O(R^{-2}), \tag{17}$$

$$R\sin\theta\log(2R/(kae))-R(\pi/2-\theta)\cos\theta=\frac{h_2\cos\theta-Q_1\sin\theta}{R}+(n+\tfrac{1}{4})\pi+O(R^{-2}) \tag{18}$$

where $Q_1=(3k^2a^2+12h_1-1)/12$. Since $0\leqq\theta\leqq\pi/2$ both of the terms appearing on the left hand side of equation (17) are positive or zero whilst the expression on the right hand side tends to zero as $R\to\infty$. Thus $\theta\to\pi/2$ as $\mathrm{R}\to\infty$. We can therefore obtain an asymptotic estimate of the solution by setting $\theta=\pi/2-\varepsilon$ where ε is small and solving for ε. This leads to the asymptotic formula

$$\theta=\frac{\pi}{2}-\frac{h_2}{R^2\log(2R/(ka))}+O[R^{-4}(\log R)^{-2}], \tag{19}$$

where R is given by the equation

$$R\log(2R/(kae))=(n+\tfrac{1}{4})\pi-\frac{Q_1}{R}+O(R^{-2}). \tag{20}$$

It follows from these equations that the real and imaginary parts of u are such that

$$s=\operatorname{Re}(u)=R\cos\theta=\frac{h_2}{R\log(2R/(ka))}[1+O\{R^{-2}(\log R)^{-1}\}], \tag{21}$$

$$t=\operatorname{Im}(u)=R\sin\theta=R[1+O\{R^{-4}(\log R)^{-2}\}] \tag{22}$$

as $R\to\infty$. It is noted that $\operatorname{Re}(u)\to 0$ as $R\to\infty$ so that the zeros move closer and closer to the imaginary axis as $R\to\infty$.

3. The Green's function. The main result established in this paper can be stated as follows:

THEOREM. *Suppose that $f(r)$ is twice continuously differentiable for $r\geqq a$, $r^{-1/2}f(r)$ and $r^{-1/2}(r^2f_{rr}+rf_r-k^2r^2f)\in L^2(a,\infty)$, where $k>0$. Then, if* $\operatorname{Im}(h)\neq 0$,

$$f(r)=-2\sum_{u=u_n}\frac{[hI_u(ka)+kaI'_u(ka)]K_u(kr)uF(u)}{(\partial/\partial u)[hK_u(ka)+kaK'_u(ka)]},\qquad r>a$$

where the summation includes all those zeros of the function $hK_u(ka)+kaK'_u(ka)$ that lie in the half plane $\operatorname{Re}(u)>0$ *and*

$$F(u)=\int_a^\infty f(r)K_u(kr)\frac{dr}{r}.$$

A proof of the above theorem can be obtained by following the method used in [5], [7] to derive a related expansion theorem. Let $f(r)$, the function to be expanded, satisfy the conditions of the theorem and define

$$r^2f_{rr}+rf_r-(k^2r^2+v^2)f=\psi,\qquad r\geqq a, \tag{23}$$

where $v>0$. The equation (23) is regarded as a nonhomogeneous equation for $f(r)$ which will be inverted by means of a suitable Green's function $G(r,\rho)$. The appropriate Green's function, which must satisfy the condition (3) at $r=a$, is given by the formulas

$$G(r,\rho)=\begin{cases}-\dfrac{[g(v)I_v(kr)-g_1(v)K_v(kr)]K_v(k\rho)}{g(v)}, & a\leqq r\leqq\rho,\\[2ex] -\dfrac{[g(v)I_v(k\rho)-g_1(v)K_v(k\rho)]K_v(kr)}{g(v)}, & a\leqq\rho\leqq r,\end{cases} \tag{24}$$

where

$$g_1(v)=hI_v(ka)+kaI'_v(ka). \tag{25}$$

Upon inverting (23) with the aid of the above Green's function and using the result (A.2) established in the Appendix to this paper it is found that

$$f(r)=\int_a^\infty \psi(\rho)G(r,\rho)\frac{d\rho}{\rho}+[af'(a)+hf(a)]\frac{K_v(kr)}{g(v)}. \tag{26}$$

To obtain the expansion formula the Green's function will now be represented as a contour integral, which will be inserted in (26). The representation in question is given by the formula

$$G(r,\rho)=\frac{1}{i\pi}\int_L\frac{[g(u)I_u(kr)-g_1(u)K_u(kr)]K_u(k\rho)u\,du}{(u^2-v^2)g(u)}. \tag{27}$$

In this formula, which is valid for all $r, \rho \geqq a$, the path L denotes a line $(c-i\infty, c+i\infty)$ parallel to the imaginary axis positioned so that all the zeros of $g(u)$ lie to the left of it. The constant c is necessarily positive. The parameter v is chosen so that $v > c$ so as to ensure that the pole at $u=v$ lies to the right of L. The equivalence of (24) and (27) may be proved by means of the calculus of residues in which the contour is closed on the right by means of a suitable sequence of semicircles which recede to infinity. In the domain Re $(u) \geqq c$ the integrand appearing in (27) is an analytic function of u except for a simple pole at the point $u=v$. When $\rho \geqq r$ the integral around the semicircle vanishes in the limit as its radius tends to infinity, so that in this case the integral along L can be evaluated by computing the residue at the pole v. When this procedure is carried out the first of the expressions quoted in (24) is obtained. To justify the above argument it is necessary to verify that the contribution from the added semicircle does vanish as the radius tends to infinity and this requires an investigation of the behavior of the integrand as $u \to \infty$. With this in view the K type Bessel functions will be expressed in terms of the I type functions by means of the relation (11). This shows that

$$(28)\qquad \begin{aligned} I_u(kr)K_u(ka)-I_u(ka)K_u(kr) &= \frac{\pi}{2\sin u\pi}[I_u(kr)I_{-u}(ka)-I_u(ka)I_{-u}(kr)] \\ &= \frac{1}{2u}[(r/a)^u-(a/r)^u][1+O(u^{-1})]. \end{aligned}$$

The equation (28) applies as $u \to \infty$ and is obtained after substituting from (12) and using the identity $\Gamma(1+u)\Gamma(1-u)\sin u\pi = u\pi$. Similarly it is found that

$$(29)\qquad I_u(kr)K'_u(ka)-I'_u(ka)K_u(kr) = -\frac{1}{2ka}\left[\left(\frac{r}{a}\right)^u+\left(\frac{a}{r}\right)^u\right][1+O(u^{-1})]$$

as $u \to \infty$. On bearing in mind the definitions (4), (25) of the functions $g(u)$, $g_1(u)$ it is found on combining the above estimates (28), (29) that

$$(30)\qquad g(u)I_u(kr)-g_1(u)K_u(kr) = -\frac{1}{2}\left[\left(\frac{r}{a}\right)^u+\left(\frac{a}{r}\right)^u\right][1+O(u^{-1})]$$

as $u \to \infty$.

Next an estimate for the function $g(u)$ will be obtained. We first employ the equation (12) to form the equation

$$hI_u(ka)+kaI'_u(ka) = \frac{(ka/2)^u}{\Gamma(u+1)}\left[h+u+\frac{(h+u+2)k^2a^2}{4(u+1)}+O(u^{-1})\right].$$

If this equation be subtracted from that obtained by changing the sign of u wherever it appears we find from (4) and (11) that

$$(31)\qquad g(u) = -\frac{\pi}{2\sin u\pi}\left[\frac{(ka/2)^u}{\Gamma(u)}-\frac{(ka/2)^{-u}}{\Gamma(-u)}\right][1+O(u^{-1})]$$

as $u \to \infty$. Finally on employing Stirling's formula (14) to estimate the Γ-functions we find the formulas

$$(32)\qquad \begin{aligned} g(u) = &-\sqrt{2\pi u}\, e^{(1/2)iu\pi+(1/4)i\pi} \\ &\cdot \sinh\left[u\log\left(\frac{2u}{kae}\right)-\frac{iu\pi}{2}-\frac{i\pi}{4}\right]\left[1+O\left(\frac{1}{u}\right)\right], \end{aligned}$$

$$K_u(k\rho)=\sqrt{2\pi/u}\,e^{(1/2)iu\pi+(1/4)i\pi}$$

(33)
$$\cdot\cosh\left[u\log\left(\frac{2u}{k\rho e}\right)-\frac{iu\pi}{2}-\frac{i\pi}{4}\right]\left[1+O\left(\frac{1}{u}\right)\right].$$

The formulas (32), (33) apply as $u\to\infty$ in the sector $0<\delta\leqq\arg u\leqq\pi/2$ and have been derived from (31) and (11) after using the relation $\sin u\pi\sim(i/2)\,e^{-iu\pi}$. If $u\to\infty$ in the sector $-\pi/2\leqq\arg u\leqq-\delta<0$ we use $\sin u\pi\sim-(i/2)\,e^{+iu\pi}$ and find the alternative formulas

$$g(u)=-\sqrt{2\pi u}\,e^{-(1/2)iu\pi-(1/4)i\pi}$$

(34)
$$\cdot\sinh\left[u\log\left(\frac{2u}{kae}\right)+\frac{iu\pi}{2}+\frac{i\pi}{4}\right]\left[1+O\left(\frac{1}{u}\right)\right],$$

$$K_u(k\rho)=\sqrt{2\pi/u}\,e^{-(1/2)iu\pi-(1/4)i\pi}$$

(35)
$$\cdot\cosh\left[u\log\left(\frac{2u}{k\rho e}\right)+\frac{iu\pi}{2}+\frac{i\pi}{4}\right]\left[1+O\left(\frac{1}{u}\right)\right].$$

The behavior of $g(u)$ as $u\to\infty$ in the sector $|\arg u|\leqq\delta$ can still be obtained from (31) provided that u is kept bounded away from the integers. It is convenient therefore to close the contour by means of a path which is made up of the two circular arcs $u=R_0\,e^{i\theta}$, $\delta\leqq|\theta|\leqq\pi/2$, connected by the straight segment $u=n+\frac{1}{2}+it$ for $|\theta|\leqq\delta$. The radius R_0 is chosen so that $R_0\cos\delta=n+\frac{1}{2}$ which ensures that a continuous path is formed, and n is taken to be a positive integer which tends to infinity. On the straight segment $\mathrm{Re}\,(u)=n+\frac{1}{2}$ and it follows that $|\sin u\pi|=\cosh\,(R\sin\theta)$ which is bounded away from zero. The formula (31) can be applied to show that

$$g(u)=-\sqrt{\pi u/2}\,e^{u\log(2u/kae)}[1+O(u^{-1})]. \tag{36}$$

This applies as $u\to\infty$ on the sequence of straight lines $\mathrm{Re}\,(u)=n+\frac{1}{2}$, $|\arg u|\leqq\delta$ as $n\to\infty$. Similarly from (11) it follows, as $u\to\infty$ on the same sequence of segments, that

$$K_u(k\rho)=\sqrt{\pi/2u}\,e^{u\log(2u/k\rho e)}[1+O(u^{-1})] \tag{37}$$

so that, by (36), (37),

$$\left|\frac{K_u(k\rho)}{g(u)}\right|=O\left[\frac{1}{R}\,e^{-R\cos\theta\log(\rho/a)}\right] \tag{38}$$

on the stated segments.

Now $|\cosh\,(A+iB)|\leqq\cosh A\leqq e^{|A|}$ and $|\sinh\,(A+iB)|\geqq|\sinh A|\geqq\frac{1}{4}\,e^{|A|}$ whenever A,B are real and $|A|\geqq\frac{1}{2}\log 2$ so that from (32), (33) and (34), (35) it follows that

$$\left|\frac{K_u(k\rho)}{g(u)}\right|\leqq\frac{1}{R}\,e^{-R\cos\theta\log(\rho/a)}\left[1+O\left(\frac{1}{R}\right)\right] \tag{39}$$

as $R\to\infty$ in the half plane $\mathrm{Re}\,(u)\geqq c>0$. Upon combining the estimates (30) and (38) or (39) as the case may be we find that the integrand appearing in (27) is

$$O[R^{-2}\,e^{-R\cos\theta\log(\rho/r)}]$$

which tends to zero as $R\to\infty$ in the stated half plane provided that $\rho\geqq r$.

If $\rho<r$ the equality of the contour integral (27) with the second of the expressions quoted in (24) for the Green's function cannot be established by an argument similar to that just used since in this case the integral over the semicircle does not tend to zero

as the radius tends to infinity. The validity of (27) when $\rho<r$ can however be established by the following alternative method which shows that the expression (27) proposed for the Green's function is indeed a symmetric function of the variables ρ, r. Upon interchanging the variables r, ρ in the equation (27) and subtracting it follows that

$$G(r,\rho)-G(\rho,r)=\frac{1}{i\pi}\int_L \frac{[I_u(kr)K_u(k\rho)-I_u(k\rho)K_u(kr)]u\,du}{u^2-v^2}. \tag{40}$$

It follows from (28), by analogy, that the cross product of Bessel functions appearing in the preceding formula is an even function of u and that

$$I_u(kr)K_u(k\rho)-I_u(k\rho)K_u(kr)=\frac{1}{2u}\left[\left(\frac{r}{\rho}\right)^u-\left(\frac{\rho}{r}\right)^u\right]\left[1+O\left(\frac{1}{u}\right)\right].$$

as $u\to\infty$. The integrand appearing in (40) is $O(u^{-2})$ as $u\to\infty$ in the strip $0\leqq \mathrm{Re}\,(u)\leqq c$ and is a regular function of u there since $v>c$. The path L may therefore be deformed parallel to itself until it coincides with the imaginary axis whereupon it is seen that the value of the integral is zero since the integrand is an odd function of u. Hence $G(r,\rho)=G(\rho,r)$ for all values of r and $\rho\geqq a$. Therefore the value of the integral appearing in (27) when $\rho<r$ can be obtained by interchanging r and ρ and evaluating the integral as before by taking the residue at the pole, a procedure which yields the second of the expressions appearing on the right hand side of equation (24).

4. The integral theorem. When the expression (27) for the Green's function is inserted into the equation (26) and the order of integration in the repeated integral reversed we find the formula

$$f(r)=\frac{1}{i\pi}\int_L \frac{[g(u)I_u(kr)-g_1(u)K_u(kr)]\Psi(u)u\,du}{(u^2-v^2)g(u)} + [af'(a)+hf(a)]\frac{K_v(kr)}{g(v)} \tag{41}$$

where

$$\Psi(u)=\int_a^\infty \psi(\rho)K_u(k\rho)\frac{d\rho}{\rho}. \tag{42}$$

To justify the above interchange of the orders of integration it is necessary to obtain a bound on the quotient $\Psi(u)/g(u)$ valid on L. A bound on the function $\Psi(u)$ can be found on applying the Schwarz inequality to (42), which gives the inequality

$$|\Psi(u)|\leqq\left\{\int_a^\infty |\psi(\rho)|^2\frac{d\rho}{\rho}\int_a^\infty |K_u(k\rho)|^2\frac{d\rho}{\rho}\right\}^{1/2}. \tag{43}$$

The Bessel function integral appearing here can be obtained by setting $y=K_u(kr)$ in equation (6) and taking the imaginary part. This yields the formula

$$2st\int_a^\infty |K_u(kr)|^2\frac{dr}{r}=-\mathrm{Im}\,[ka\overline{K_u(ka)}K'_u(ka)]. \tag{44}$$

If the expression (33) or (35), as the case may be, is substituted into (44) it is found that

$$\int_a^\infty |K_u(kr)|^2\frac{dr}{r}=\frac{\pi e^{-\pi|R\sin\theta|}}{2R^2\sin\theta\cos\theta}[\sin\theta\sinh 2A+\cos\theta\sin 2B]\left[1+O\left(\frac{1}{R}\right)\right] \tag{45}$$

where $A+iB$ denotes the argument in the hyperbolic functions appearing in (33) and (35). Thus on setting $u=R\,e^{i\theta}$ and separating real and imaginary parts it is found that

$$(46)\quad A=\begin{cases} R\cos\theta\log(2R/(kae))+(\pi/2-\theta)R\sin\theta, & 0<\delta\leqq\theta\leqq\pi/2,\\ R\cos\theta\log(2R/(kae))-(\pi/2+\theta)R\sin\theta, & -\pi/2\leqq\theta-\delta<0,\end{cases}$$

$$(47)\quad B=\begin{cases} R\sin\theta\log(2R/(kae))-(\pi/2-\theta)R\cos\theta-\pi/4, & 0<\delta\leqq\theta\leqq\pi/2,\\ R\sin\theta\log(2R/(kae))+(\pi/2+\theta)R\cos\theta+\pi/4, & -\pi/2\leqq\theta\leqq-\delta<0.\end{cases}$$

The behavior of $g(u)$ can be deduced from (32) and (34) which show that

$$(48)\qquad |g(u)|=\sqrt{2\pi R}\,e^{-(\pi/2)|R\sin\theta|}(\cosh^2 A-\cos^2 B)^{1/2}\left[1+O\left(\frac{1}{R}\right)\right]$$

as $R\to\infty$ in $0<\delta\leqq|\theta|\leqq\pi/2$.

On the path L, $R\to\infty$ whilst $R\cos\theta=c$, a constant. Therefore $\theta=\pm(\pi/2-c/R)+O(R^{-3})$ so that the quantity A defined by (46) is such that $A\to c\log(2R/kae)+c$ which tends to infinity with R. It follows from (43), (45) that

$$(49)\qquad \Psi(u)=O\left[\frac{e^{-(\pi/2)|R\sin\theta|}}{|R\sin\theta|^{1/2}}(\sinh 2A)^{1/2}\right].$$

Since $\sinh 2A/(\cosh^2 A-\cos^2 B)\to 2$ as $A\to\infty$ then $\Psi(u)/g(u)=O(u^{-1})$ as $u\to\infty$ on L and therefore by (30) the integrand appearing in (41) is $O(u^{-2})$ as $u\to\infty$ on L. Hence the integral along L is absolutely convergent and so the interchange in the order of integration is justified.

To complete the derivation of the expansion formula quoted in the theorem of § 3 it is now necessary to express the transform $\Psi(u)$ in terms of the basic transform $F(u)$ defined by (1). To do this the equation (23) is multiplied by $r^{-1}K_u(kr)$ and integrated for $a\leqq r<\infty$, the term involving $(r^2 f_{rr}+rf_r)$ being integrated twice by parts. If the resulting term involving the limit as $r\to\infty$ be eliminated by appealing to equation (A.1), established in the Appendix to this paper, it is found that

$$(50)\qquad \Psi(u)=(u^2-v^2)F(u)-[af'(a)+hf(a)]K_u(ka)+g(u)f(a).$$

Upon inserting this result in (41) we find the equation

$$(51)\qquad f(r)=\frac{1}{i\pi}\int_L\frac{[g(u)I_u(kr)-g_1(u)K_u(kr)]F(u)u\,du}{g(u)} +Bf(a)-A[af'(a)+hf(a)]$$

where

$$A=\frac{K_v(kr)}{g(v)}-\frac{1}{i\pi}\int_L\frac{[g(u)I_u(kr)-g_1(u)K_u(kr)]K_u(ka)u\,du}{(u^2-v^2)g(u)},$$

$$B=\frac{1}{i\pi}\int_L\frac{[g(u)I_u(kr)-g_1(u)K_u(kr)]u\,du}{u^2-v^2}.$$

It will now be shown that $A=B=0$.

First we note that the integral appearing in the expression for A is the same as that appearing in (27) with ρ set equal to a therein and it is therefore equal to $G(r,a)$. On calculating $G(r,a)$ by setting $\rho=a$ in the second of the expressions given in (24) and utilizing the Wronskian identity $I_u(x)K'_u(x)-K_u(x)I'_u(x)=-x^{-1}$ it follows that $G(r,a)=K_v(kr)/g(v)$ so that $A=0$ as required. To show that $B=0$ we first observe that the combination of Bessel functions present in the integral for B reduces to an even function of the variable u, as can readily be verified on expressing K_u in terms of I_u and I_{-u} by means of (11). We also note by (30) that the path L may be deformed onto the imaginary axis whereupon it is seen that the integral is equal to zero since the integrand is an odd function of u. The equation (51) then reduces to the integral formula

$$f(r)=\frac{1}{i\pi}\int_L \frac{[g(u)I_u(kr)-g_1(u)K_u(kr)]F(u)u\,du}{g(u)}. \tag{52}$$

This formula, in which the path L is situated so that all of the zeros of $g(u)$ lie to the left of it, is actually the formula of inversion associated with the integral transform defined by the equation (1), and for some purposes it is preferable to the actual expansion in eigenfunctions stated in the theorem of § 3. To obtain the series form of the expansion the path L appearing in (52) is deformed onto the imaginary axis. Since the functions $g(u)$ and $F(u)$ are, like K_u itself, even in u and, as noted above, the remaining Bessel function terms also form an even function of u consequently the integrand in (52) is an odd function of u so that the value of the integral when taken along the entire imaginary axis is zero. Upon evaluating the residues at the poles situated between L and the imaginary axis we find the formula

$$f(r)=-2\sum_{u=u_n}\frac{ug_1(u)K_u(kr)F(u)}{g'(u)}. \tag{53}$$

On recalling the definition (25) of $g_1(u)$ it is seen that the above series is the same as that stated in the theorem.

To justify (53) it is necessary to apply Cauchy's theorem to a sequence of rectangles formed by connecting L to the imaginary axis by means of the two straight lines $\mathrm{Im}\,(u)=\pm t_n$, $0\leqq \mathrm{Re}\,(u)\leqq c$. The ordinate t_n is chosen so that the top and bottom sides of the rectangle pass midway between the zeros of the function $g(u)$. By (20) this will be achieved if we select $t_n \log(2t_n/(kae))=(n+\frac{3}{4})\pi$, since $\mathrm{Im}\,(u)=R+O(R^{-1})$ as $u\to\infty$ in the strip $0\leqq \mathrm{Re}\,(u)\leqq c$. Also as $u\to\infty$ in this strip, $\cos\theta=O(R^{-1})$ and $\pi/2-|\theta|=O(R^{-1})$ so that the angle B defined by (46), (47) can be written as

$$|B|=|t|\log|2t/(kae)|-(\pi/2-|\theta|)R\cos\theta-\pi/4=(n+\tfrac{1}{2})\pi+O(R^{-1}).$$

It follows that $\cos B=O(R^{-1})$ and therefore, by (48),

$$|g(u)|=\sqrt{2\pi R}\,e^{-(\pi/2)|R\sin\theta|}(\cosh A)[1+O(R^{-1})] \tag{54}$$

valid on the lines $0\leqq \mathrm{Re}\,(u)\leqq c$, $\mathrm{Im}\,(u)=\pm t_n$.

A bound on the function $F(u)$ valid in the strip can be obtained from equation (50), the asymptotic behavior of the function $\Psi(u)$ being deduced from (43) and (45). In applying the formula (45) care must be taken in the vicinity of the imaginary axis since the factor $\cos\theta$ is zero on that axis. This difficulty can be overcome by noting that $\pi/2-|\theta|\leqq(\pi/2)\cos\theta$ for $|\theta|\leqq\pi/2$ so that the quantity A defined in (46), (47) is such that

$$0\leqq A\leqq R\cos\theta[\log(2R/(kae))+(\pi/2)|\sin\theta|].$$

Now $\sinh bx\leqq b\sinh x$ whenever $0\leqq b\leqq 1$ and $x\geqq 0$ so that

$$\sinh 2A\leqq R\cos\theta\sinh[2\log(2R/(kae))+\pi|\sin\theta|]$$

provided that $|R\cos\theta|\leqq 1$. It follows from (43), (45) and the above inequality that

$$(55)\quad \begin{aligned}\Psi(u)&=O\left[|R\sin\theta|^{-1/2}e^{-(\pi/2)|R\sin\theta|}\left\{1+\sinh\left(2\log\left(\frac{2R}{kae}\right)+\pi\right)\right\}^{1/2}\right]\\&=O\left[\left|\frac{R}{\sin\theta}\right|^{1/2}e^{-(\pi/2)|R\sin\theta|}\right]\end{aligned}$$

as $u\to\infty$ in the strip $0\leqq \mathrm{Re}\,(u)\leqq 1$. On combining (54) and (55) it follows that Ψ/g is $O(\operatorname{sech} A)$ which is bounded as $u\to\infty$ on the sequence of sides $\mathrm{Im}\,(u)=\pm t_n$.

If $c>1$ it is also necessary to consider the additional region $1\leqq\mathrm{Re}\,(u)\leqq c$, however since now $R\cos\theta\geqq 1$ this presents no difficulty. The quantity A now tends to infinity as $R\to\infty$ and it is clear from the formulas (43), (45) that the bound (49) holds. Since (54) is also valid it follows that $\Psi/g=O(u^{-1})$ as $u\to\infty$ on the sequence of sides $\mathrm{Im}\,(u)=\pm t_n$ in the strip $1\leqq\mathrm{Re}\,(u)\leqq c$.

The behavior of the quotient $F(u)/g(u)$ can now be deduced from that of Ψ/g by dividing the equation (50) by $g(u)$, since $K_u(ka)=O(R^{c-1/2}e^{-(\pi/2)R})$ as $u\to\infty$ on the stated sequence of lines t_n. The cross product of Bessel functions appearing in (52) is by (30) bounded as $u\to\infty$ in the strip, since $\mathrm{Re}\,(u)$ is bounded there, so that the integrand appearing in (52) is $O(u^{-1})$. This verifies that the integrals along the sides $\mathrm{Im}\,(u)=\pm t_n$ tend to zero in the limit as $u\to\infty$ so that the formula (53) is established.

Appendix. In this Appendix the conditions on $f(r)$ imposed in the expansion theorem are used to derive an asymptotic bound on the derivative $f'(r)$. The bound in question, which is necessary to complete the derivation of equations (26) and (50), can be obtained by forming the equations

$$\begin{aligned}Rf'(R)-af'(a)&=\int_a^R (rf_{rr}+f_r)\,dr\\&=\int_a^R r^{-1/2}(r^2f_{rr}+rf_r-k^2r^2f)r^{-1/2}\,dr+k^2\int_a^R (r^{-1/2}f)r^{3/2}\,dr\\&=I_1(R)+k^2I_2(R).\end{aligned}$$

Now on applying the Schwarz inequality to I_1 and I_2 and using the integrability conditions stated in the theorem it is found that

$$|I_2|\leqq\left(\int_a^R|r^{-1/2}f|^2\,dr\cdot\int_a^R r^3\,dr\right)^{1/2}=O(R^2),$$

$$|I_1|\leqq\left(\int_a^R|r^{-1/2}[r^2f_{rr}+rf_r-k^2r^2f]|^2\,dr\cdot\int_a^R r^{-1}\,dr\right)^{1/2}=O[(\log R)^{1/2}].$$

It follows from these results that $f'(R)=O(R)$ and therefore by integration that $f(R)=O(R^2)$ as $R\to\infty$. Since $K_u(kr)$ and $K'_u(kr)$ are $O(r^{-1/2}e^{-kr})$ as $r\to\infty$ then

$$(\mathrm{A}.1)\qquad \lim_{r\to\infty} r[f'(r)K_u(kr)-kf(r)K'_u(kr)]=0$$

and therefore, by the second expression of (24),

$$(\mathrm{A}.2)\qquad \lim_{r\to\infty} r[f'(r)G(r,\rho)-f(r)G_r(r,\rho)]=0.$$

REFERENCES

[1] D. S. COHEN, *Eigenfunction expansions and nonself adjoint boundary value problems*, Comm. Pure Appl. Math., 17 (1964), pp. 23–34.

[2] D. NAYLOR, *An eigenvalue problem in cylindrical harmonics*, J. Math. and Phys., 44 (1965), pp. 391–402.

[3] ———, *On some eigenvalue problems involving modified Bessel functions*, Ibid., 45 (1966), pp. 95–103.

[4] ———, *An eigenfunction expansion associated with a condition of radiation*, Proc. Cambridge Philos. Soc., 67 (1970), pp. 107–121.

[5] ———, *On an integral transform associated with a condition of radiation*, Ibid., 71 (1972), pp. 369–379.

[6] D. NAYLOR, F. C. CHOO AND D. W. BARCLAY, *On an eigenfunction expansion associated with a condition of radiation*, part 2, Ibid., 74 (1973), pp. 485–496.

[7] D. NAYLOR AND D. W. BARCLAY, *On an expansion problem occurring in the theory of diffraction*, this Journal, 6 (1975), pp. 199–207.

[8] E. PFLUMM, *Expansion problems arising from the Watson transformation.* Report BR 35, Division of Electromagnetic Research, New York University, New York, 1960.

[9] E. C. TITCHMARSH, *Eigenfunction Expansions Associated with Second Order Differential Equations*, part 1, Oxford University Press, London, 1962.

[10] G. N. WATSON, *Theory of Bessel Functions*, Cambridge University Press, London, 1958.

SIAM J. MATH. ANAL.
Vol. 9, No. 6, December 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0001 $01.00/0

A MULTI-TIME SCALE METHOD IN ALMOST-PERIODIC STABILITY PROBLEMS*

JON H. DAVIS†

Abstract. This paper considers the problem of stability of a class of almost-periodic ordinary differential equations. Necessary and sufficient stability conditions are derived by use of a method of "multiple time scales" suggested by similar techniques in perturbation theory. The conditions obtained involve a spectral analysis of a compact operator determined by the original system coefficients, and in principle are computable to any desired degree of accuracy by finite dimensional methods.

1. Introduction. This paper is concerned with differential equations of the form

$$p(D)x(t)+k(t)q(D)x(t)=u(t), \qquad t\geqq 0, \tag{1}$$

arising in the study of the stability properties of feedback control systems. It is known (see [5], for example) that the input-output stability of (1) in the $L_2(0, \infty)$ sense is equivalent to the invertibility of the so-called minimal operator associated with (1) on $L_2(0, \infty)$.

There exists a considerable amount of literature devoted to the derivation of sufficient conditions for either stability (invertibility) or instability (noninvertability) of the system (1) under various hypotheses on the character of the "time-varying gain" function $k(\cdot)$. (The articles [1], [2], [3], [4], [5] contain various references to problems of this sort). The problem of obtaining necessary and sufficient conditions for stability of course still remains—in fact it is of interest to ask for which classes of systems is it possible to give conditions which are computable in terms of the given system parameters. In the control theoretic context, the case of constant coefficients is readily handled by the Nyquist criterion. The use of Floquet's Theorem for the case of periodic coefficients permits progress in obtaining computable conditions; one approach is through calculation of the characteristic exponents of the system. A method much more closely related to the results of this paper is in [5], where a generalization of the Nyquist criterion to the case of periodic systems was obtained by use of what is essentially a vector-valued transform technique.

This paper reports an "extension" of the above method to the case of a class of differential equations with almost-periodic coefficients. The fact that there exists no complete analog of Floquet's result for the case of almost-periodic coefficients (in the sense that it is not in general true that an almost-periodic system may be reduced to a constant-coefficient system by means of an almost-periodic Lyapunov transformation, see [9], [10], for example), guarantees that the almost-periodic problem will be more complicated than the periodic case. However, it is possible to combine a method of multiple time scales related to certain perturbation-theoretic methods [8] with the transform method of [5] in order to derive necessary and sufficient conditions for the stability of system (1). These results reduce the question of invertibility of the given differential operator on a half-axis to that of the invertibility of a related partial differential operator on a half-space. The spectrum of the partial differential operator may be calculated in terms of the locus of the spectrum of a *compact* operator depending on a "transform variable". This compact operator may be explicitly determined from the parameters of the system (1), and so in principle the stability

* Received by the editors, May 13, 1976, and in final revised form January 31, 1977.
† Department of Mathematics, Queen's University, Kingston, Ontario, Canada K7L 3N6.

criterion obtained below is computable to any desired degree of accuracy by finite dimensional methods.

The basic methods used here and the results obtained are relatively simple to state, and are contained in § 2 below.

Proofs of these results are reserved for § 3.

2. An outline of the method and results. The results of this paper arose out of an attempt to extend known methods of spectral computation for the singular integral equation

$$e(t)+k(t)\int_0^t g(t-s)e(s)\,ds=u(t),\qquad t\geqq 0, \tag{2}$$

beyond the cases of constant and periodic $k(\cdot)$. In those cases the known results have an interpretation as a "factoring" of the spectrum of the operator through a transform variable (dual to the additive semi-group associated with the original half-axis problem). This identifies the spectrum of the operator in (2) as the range of the Laplace transform in the case that $k(\cdot)$ is constant, and essentially as the range of the spectrum of an associated compact operator involving a transform variable in the case of periodic $k(\cdot)$ [5]. The work in this paper was motivated by a desire to obtain results having a similar degree of computational potential for the case of an almost periodic gain $k(\cdot)$. This paper assumes that the function $g(\cdot)$ in (2) has a rational Laplace transform

$$\mathscr{L}\{g(\cdot)\}(s)=q(s)/p(s).$$

With this assumption, (2) has an interpretation as the integral equation for the "return difference" of a feedback control system governed by a system of ordinary differential equations. If $[\mathbf{A},\mathbf{b},\mathbf{c}']$ is a minimal realization of the transfer function

$$q(s)/p(s)=\mathbf{c}'(\mathbf{I}s-\mathbf{A})^{-1}\mathbf{b},$$

then the feedback control system under consideration is governed by

$$\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}+\mathbf{b}(u(t)-k(t)\mathbf{c}'\mathbf{x}).$$

With the return difference function defined by

$$e(t)=u(t)-k(t)\mathbf{c}'\mathbf{x}(t)$$

the equation (2) follows easily in the case of zero initial conditions for the ordinary differential equation. Assuming without loss of generality that the realization $[\mathbf{A},\mathbf{b},\mathbf{c}']$ is in standard controllable form, the vector differential equation reduces to (in the usual notation)

$$p(D)x(t)+k(t)q(D)x(t)=u(t),\qquad t\geqq 0, \tag{3}$$

with the boundary conditions $x(0)=Dx(0)=D^{(\deg p-1)}x(0)=0$.

The equivalence of the two system descriptions from the point of view of input-output stability follows readily in the case of a stable $\mathbf{A}$ matrix. In the case of an unstable $\mathbf{A}$, stability equivalence relies on uniform observability of the system with feedback.

More detailed descriptions of the relationship of the two system models (2) and (3) may be found, for example, in [4, §§ 15–19, 33–35], and [6, Chap. 6]. We also restrict attention to the case in which $k(\cdot)$ contains only finitely many "fundamental frequencies", that is

$$k(t)=\sum_{i=1}^{n}k_i(\omega_i t),$$

with $k_i(t)=k_i(t+1)$, and $\{\omega_i\}_{i=1}^n$ linearly independent over the rationals. In fact, we shall write the formulae and present arguments below for the case $n=2$—the case $n>2$ essentially requires only more complicated notation, and no changes in the basic methods or arguments involved.

If one attempts formally to find Floquet-type solutions (i.e. solutions of the form of an almost-periodic function multiplied by an exponential function) to the homogeneous equation (3) (assuming now, of course, nonzero initial conditions), then it soon becomes apparent that a major source of difficulty in carrying out the formal calculation is the denseness of the set of points $\{j\omega_1+k\omega_2\}_{j,k\in Z}$ in R^1.

We avoid this problem by "decoupling" the effects of the separate feedback frequencies. This is accomplished by introducing additional independent variables into the problem. Our methods are suggested by the method of "multiple time scales" in perturbation theory, ([8], for example) in that we seek to avoid difficulties caused by the presences of the two "natural" times scales ($1/\omega_1$ and $1/\omega_2$) in the differential equation

$$\left[p\left(\frac{d}{dt}\right)+\{k_1(\omega_1 t)+k_2(\omega_2 t)\}q\left(\frac{d}{dt}\right)\right]x(t)=u(t),\qquad t\geqq 0,$$

by introducing *two* time scales $\tau=\omega_1 t$ and $\sigma=\omega_2 t$. In terms of these new variables the problem formally becomes

$$\left[p\left(\omega_1\frac{\partial}{\partial\tau}+\omega_2\frac{\partial}{\partial\sigma}\right)+\{k_1(\tau)+k_2(\sigma)\}q\left(\omega_1\frac{\partial}{\partial\tau}+\omega_2\frac{\partial}{\partial\sigma}\right)\right]\tilde{x}(\tau,\sigma)=\tilde{u}(\tau,\sigma),\qquad \tau,\sigma\geqq 0. \tag{4}$$

This formal procedure finds some success in perturbation problems arising in certain applied mathematical contexts in which the forcing function $u(\cdot)$ is *explicitly known*, and a *specific* solution $x(\cdot)$ is sought. In the general problem, it seems difficult to find a suitable mapping from the space of original forcing functions $\{u(\cdot)\}$ to the two variable functions $\{\tilde{u}(\cdot,\cdot)\}$ since the formal requirement $\tilde{u}(\omega_1 t,\omega_2 t)=u(t)$ is far from a condition uniquely determining $\tilde{u}(\cdot,\cdot)$. Of course, in a specific problem with *given* $u(\cdot)$, this ambiguity may be exploited in order to facilitate computations. A survey of multiple time scale methods in perturbation problems appears in [15].

In the present problem, however, our interest is essentially in the solvability of (3) for arbitrary $u(\cdot)\in L_2[0,\infty)$. Specifically, we wish to establish connections between invertibility of the original ordinary differentential operator, and invertibility of the partial differential operator defined by the formal expression (4) obtained above.

Because of what seem to us to be inherent difficulties involved in attempting to formalize the process of obtaining the mapping $u(\cdot)\to\tilde{u}(\tau,\sigma)$, we avoid these considerations entirely. Instead, we justify consideration of the formal differential operator (4) by use of the following devices. We imbed the original differential operator

$$L=p(D)+\{k_1(\omega_1 t)+k_2(\omega_2 t)\}q(D) \tag{5}$$

acting in the space $L_2(R_1^+)$ in the family of operators

$$L_s=p(D)+\{k_1(\omega_1 t+\omega_2 s)+k_2(\omega_2 t+\omega_1 s)\}q(D) \tag{6}$$

depending on the parameter s, acting in the original space $L_2(R_1^+)$.

We show then that L has a bounded inverse if and only if L_s has an inverse, bounded uniformly in s, and further, that this is so if and only if the partial differential operator

$$\mathscr{L}=p\left(\frac{\partial}{\partial t}\right)+\{k_1(\omega_1 t+\omega_2 s)+k_2(\omega_2 t+\omega_1 s)\}q\left(\frac{\partial}{\partial t}\right) \tag{7}$$

acting in the space L_2 of a suitably chosen half-space of R^2 has a bounded inverse.

By a change of independent variable this is equivalent to the invertibility of

$$(P+KQ)=p\left(\omega_1\frac{\partial}{\partial\tau}+\omega_2\frac{\partial}{\partial\sigma}\right)+\{k_1(\tau)+k_2(\sigma)\}q\left(\omega_1\frac{\partial}{\partial\tau}+\omega_2\frac{\partial}{\partial\sigma}\right) \tag{8}$$

on $L_2(R_2^+)$. Since this last operator has periodic coefficients, it may be analyzed by a transform method; however, it is not the case that the resulting "transformed operator" has a compact resolvent, as is desired from a computational point of view. The intuitive reason for this is that the partial differential operator has no "dispersion" in the direction of the "artificial" additional variable s.

We introduce a factor of an operator $r(\partial/\partial s)$, and consider instead the operator

$$\begin{aligned} R(P+KQ)=r\left(\omega_2\frac{\partial}{\partial\tau}+\omega_1\frac{\partial}{\partial\sigma}\right)\Big[p\left(\omega_1\frac{\partial}{\partial\tau}+\omega_2\frac{\partial}{\partial\sigma}\right)+\{k_1(\tau)+k_2(\sigma)\} \\ \cdot q\left(\omega_1\frac{\partial}{\partial\tau}+\omega_2\frac{\partial}{\partial\sigma}\right)\Big]. \end{aligned} \tag{9}$$

The original operator (4) is invertible if and only if (8) is invertible on $L_2(R_2^+)$. Further, a transform reduces the question of the invertibility of (8) to that of an operator depending on the transform variable, and having a compact resolvent. This result may be stated as:

THEOREM 1. *Let* $k(t)=k_1(\omega_1 t)+k_2(\omega_2 t)$, $k_i(t+1)=k_i(t)\in C^\infty[0,\infty)$, *and consider the minimal differential operator* L *on* $L_2[0,\infty)$ *defined by the expression*

$$Lx(t)=\left[p\left(\frac{d}{dt}\right)+k(t)q\left(\frac{d}{dt}\right)\right]x(t)$$

on its domain. Let $r(\cdot)$ *be an arbitrary polynomial whose zeroes all have negative real part. Then the differential operator defined by the expression*

$$\begin{aligned} &R(P+KQ)(s_1,s_2)\hat{x}(\tau,\sigma) \\ &=r\left(\omega_2\frac{\partial}{\partial\tau}+\omega_1\frac{\partial}{\partial\sigma}+\omega_2 s_1+\omega_1 s_2\right)\cdot\Big[p\left(\omega_1\frac{\partial}{\partial\tau}+\omega_2\frac{\partial}{\partial\sigma}+\omega_1 s_1+\omega_2 s_2\right) \\ &\quad+\{k_1(\tau)+k_2(\sigma)\}q\left(\omega_1\frac{\partial}{\partial\tau}+\omega_2\frac{\partial}{\partial\sigma}+\omega_1 s_1+\omega_2 s_2\right)\Big]\hat{x}(\tau,\sigma), \end{aligned}$$

subject to the boundary conditions

$$\hat{x}(\tau+1,\sigma)=\hat{x}(\tau,\sigma+1)=\hat{x}(\tau,\sigma),$$

has a compact resolvent

$$\mathcal{R}(\lambda;R(P+KQ)(s_1,s_2)),$$

for $\operatorname{Re} s_1$, $\operatorname{Re} s_2$ *sufficiently large; further, the operator* L *has a bounded inverse on* $L_2(0,\infty)$ *if and only if* $\mathcal{R}(0;R(P+KQ)(s_1,s_2))$ *is analytic for all* $\operatorname{Re} s_1$, $\operatorname{Re} s_2\geqq 0$.

The above result appears to involve a multi-dimensional transform variable—however, in view of the fact that the "additional" variables were somewhat arbitrarily introduced into the problem, it might be suspected that the "extra" transform variables are superfluous. This is in fact the case.

THEOREM 2. *Consider the operator L and polynomial $r(\cdot)$ as in Theorem 1, and the differential operator defined on its domain* $\subset L_2[(0,1)\times(0,1)]$ *by*

$$R(P+KQ)(s)\hat{x}(\tau,\sigma)=r\left(\omega_2\frac{\partial}{\partial\tau}+\omega_1\frac{\partial}{\partial\sigma}\right)\left[p\left(\omega_1\frac{\partial}{\partial\tau}+\omega_2\frac{\partial}{\partial\sigma}+s\right)\right.$$
$$\left.+\{k_1(\tau)+k_2(\sigma)\}q\left(\omega_1\frac{\partial}{\partial\tau}+\omega_2\frac{\partial}{\partial\sigma}+s\right)\right]\hat{x}(\tau,\sigma),$$

subject to boundary conditions

$$\hat{x}(\tau+1,\sigma)=\hat{x}(\tau,\sigma+1)=\hat{x}(\tau,\sigma).$$

Then $[R(P+KQ)(s)]^{-1}$ *exists and is compact for sufficiently large* Re s; *L has a bounded inverse on* $L_2(0,\infty)$ *if and only if* $[R(P+KQ)(s)]^{-1}$ *exists for all* Re $s\geqq 0$, *i.e. the locus of the eigenvalues of* $R(P+KQ)(s)$ *for* Re $s\geqq 0$ *remains bounded away from the origin.*

3. Proof of main results. We consider the differential operator L of degree n, defined on the domain $\mathscr{D}(L)=\{x(\cdot)|x(\cdot), dx/dt(\cdot),\cdots,(dx^{(n-1)})/(dt^{(n-1)})(\cdot)$ all absolutely continuous, $(d^nx/dt^n)(\cdot)\in L_2(0,\infty)$, $x(0)=(dx/dt)(0)=\cdots=(dx^{(n-1)}/dt^{(n-1)})(0)=0\}\subset L_2(0,\infty)$ by

$$Lx(t)=\left[p\left(\frac{d}{dt}\right)+\{k_1(\omega_1t)+k_2(\omega_2t)\}q\left(\frac{d}{dt}\right)\right]x(t), \tag{10}$$

under the assumptions that
(i) $\deg p=n>\deg q$.
(ii) $k_i(\cdot)\in C^\infty[0,\infty)$, and $k_i(t+1)=k_i(t)$.
(iii) $\{\omega_1,\omega_2\}$ are independent over the rationals.

That is, we condsider the minimal closed operator associated with the formal differential expression (10) [7].

In addition to the operator L defined above, we consider the family of operators L_s defined by the conditions $\mathscr{D}(L_s)=\mathscr{D}(L)$, and

$$L_sx(t)=\left[p\left(\frac{d}{dt}\right)+[k_1(\omega_1t+\omega_2s)+k_2(\omega_2t+\omega_1s)]q\left(\frac{d}{dt}\right)\right]x(t). \tag{11}$$

LEMMA 1. *L has a bounded inverse on* $L_2(0,\infty)$ *if and only if* L_2 *is invertible on* $L_2(0,\infty)$ *for all* $s\in R^1$, *and* $\|L_s^{-1}\|\leqq M$, *uniformly in s.*

Proof. Clearly the existence of $L_s^{-1}\forall s$ implies L^{-1} exist simply by setting $s=0$. To show that the existence of L^{-1} implies that of L_s^{-1}, consider the differential equation $Lx(t)=u(t)$, i.e.

$$\left[p\left(\frac{d}{dt}\right)+\{k_1(\omega_1t)+k_2(\omega_2t)\}q\left(\frac{d}{dt}\right)\right]x(t)=u(t),$$

and restrict the forcing function $u(\cdot)$ to be zero almost everywhere on $[0,T)$. Then defining $t-T=\tau$, $u_T(\tau)=u(\tau+T)$, $x_T(\tau)=x(\tau+T)$ for $\tau\geqq 0$, we have

$$\begin{aligned}L_Tx_T(\tau)&=\left[p\left(\frac{d}{d\tau}\right)+\{k_1(\omega_1(\tau+T))+k_2(\omega_2(\tau+T))\}q\left(\frac{d}{d\tau}\right)\right]x_T(\tau)\\&=u_T(\tau),\qquad \tau\geqq 0,\end{aligned} \tag{12}$$

and by the hypothesis of the existence of L^{-1}, we see that $\|L_T^{-1}\|\leqq\|L^{-1}\|$, since

$$\|x_T(\cdot)\|\leqq\|L^{-1}\|\cdot\|u_T(\cdot)\|.$$

Using the hypothesis that $k_i(\cdot)$ are periodic of period 1, and the observation that the mapping

$$\varphi: T \to \begin{bmatrix} \omega_1 T \\ \omega_2 T \end{bmatrix} (\text{mod } 1),$$

is dense in $[0, 1] \times [0, 1]$ for ω_1/ω_2 irrational we conclude that with

$$L\alpha\beta\, x(\tau) = \left[p\left(\frac{d}{d\tau}\right) + \{k_1(\omega_1 t + \alpha) + k_2(\omega_2 \tau + \beta)\} q\left(\frac{d}{d\tau}\right) \right] x(\tau),$$

we have

$$\|L\alpha\beta^{-1}\| \leqq \|L^{-1}\|$$

for a dense set of $(\alpha, \beta) \in [0, 1] \times [0, 1]$.

Perturbing $L\alpha\beta$ to

$$L\alpha\beta, \varepsilon\, x(\tau) = \left[p\left(\frac{d}{d\tau}\right) + \{k_1(\omega_1 \tau + \alpha + \varepsilon) + k_2(\omega_2 \tau + \beta + \varepsilon)\} q\left(\frac{d}{d\tau}\right) \right] x(\tau),$$

we note that the hypotheses on p, q and $k_i(\cdot)$ ensure that $L\alpha\beta, \varepsilon$ is a relatively-bounded perturbation of $L\alpha\beta$. Using standard perturbation arguments (e.g. [7, Thm. V.3.6]) shows that there exists M such that

$$\|L\alpha\beta, \varepsilon\| \leqq M,$$

for all ε sufficiently small. This makes $L\alpha\beta$ uniformly invertible for $(\alpha, \beta) \in [0, 1] \times [0, 1]$, and the conclusion of the lemma now follows immediately.

COROLLARY. *We note that the above argument shows that L_s^{-1} exists, with $\|L_s^{-1}\|$ uniformly bounded in s if and only if $L_{s_0}^{-1}$ exists for at least one $s_0 \in R$.*

We denote by Ω the subset of R^2,

$$\Omega = \{(t, s) | \omega_1 t + \omega_2 s \geqq 0 \text{ and } \omega_2 t + \omega_1 s \geqq 0\},$$

and denote by Σ the boundary of Ω. $\Sigma = \{(t, s) | (t, s) \in \Omega,\ \omega_1 t + \omega_2 s = 0 \text{ or } \omega_2 t + \omega_1 s = 0\}$. (See Fig. 1).

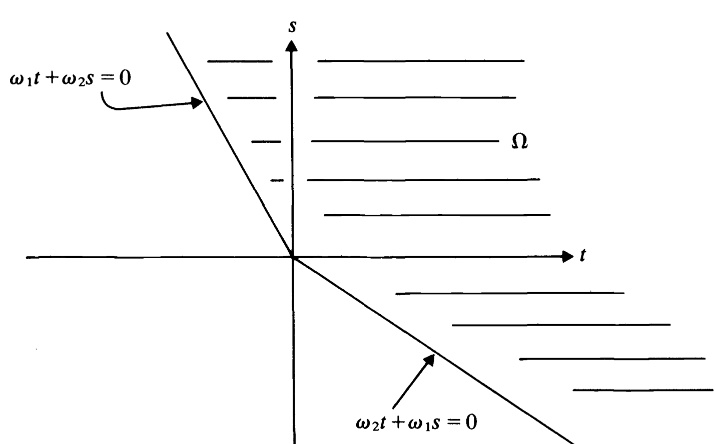

FIG. 1

We let $\mathscr{L}$ denote the linear partial differential operator defined on $L_2(\Omega)$ by the conditions $\mathscr{D}(\mathscr{L})=\{x(\cdot,\cdot)\in L_2(\Omega)$ such that the distributional derivatives $\partial x/\partial t$, $\partial^2 x/\partial t^2, \cdots \partial^n x/\partial t^n \in L_2(\Omega)$, and $x|_\Sigma = \partial x/\partial t|_\Sigma = \cdots \partial^{n-1}x/\partial t^{n-1}|_\Sigma = 0\}$, and

$$\mathscr{L}x(t,s)=\left[p\left(\frac{\partial}{\partial t}\right)+\{k_1(\omega_1 t+\omega_2 s)+k_2(\omega_2 t+\omega_1 s)\}q\left(\frac{\partial}{\partial t}\right)\right]x(t,s).$$

From the above definition of $\mathscr{L}$ it follows that $\mathscr{L}$ is a closed operator, and that the fundamental solution (Green's function) associated with the partial differential equation

$$\mathscr{L}x(t,s)=u(t,s), \qquad (t,s)\in\Omega,$$

is given by

$$g(t,\tau;s)\delta(s-\sigma)$$

with $g(\cdot,\cdot;s)$ the solution of the ordinary differential equation

$$\left[p\left(\frac{d}{dt}\right)+\{k_1(\omega_1 t+\omega_2 s)+k_2(\omega_2 t+\omega_1 s)\}q\left(\frac{d}{dt}\right)\right]g(t,\tau;s)=\delta(t-\tau), \qquad t,\tau\in\Omega,$$

subject to the boundary conditions

$$g|_\Sigma=\frac{dg}{dt}|_\Sigma=\cdots=\frac{d^{n-1}g}{dt^{n-1}}|_\Sigma=0.$$

With this notation, the unique solution to the partial differential equation

$$\mathscr{L}x(t,s)=u(t,s), \qquad (t,s)\in\Omega,$$

subject to the boundary conditions included in the definition of $\mathscr{L}$, and for $u(\cdot,\cdot)\in C_0^\infty(\Omega)$ is given by

$$x(t,s)=\int_\Sigma^t g(t,\tau;s)u(\tau,s)d\tau.$$

In the above integral, Σ represents the t-intercept of the boundary of the region Ω. From Fig. 1, we see that $\Sigma=-(\omega_2/\omega_1)s$ for $s\geqq 0$, and $\Sigma=-(\omega_1/\omega_2)s$ for $s<0$. (The indicated axis orientation corresponds to the case of $\omega_2>\omega_1$).

LEMMA 2. *Let the operator L be defined as in Lemma 1. Then L has a bounded inverse on $L_2(0,\infty)$ if and only if $\mathscr{L}$ has a bounded inverse on $L_2(\Omega)$.*

Proof. Suppose first that L^{-1} exists. Then if $u(\cdot,\cdot)\in L^2(\Omega)$, define (for almost all s)

$$x(t,s)=\int_\Sigma^t g(t,\tau;s)u(\tau,s)d\tau.$$

From Lemma 1 it follows that

$$\int_\Sigma^\infty |x(t,s)|^2 dt \leqq M\int_\Sigma^\infty |u(\tau,s)|^2 d\tau$$

with M independent of s. Hence

$$\iint_\Omega |x(t,s)|^2\,dt\,ds \leqq M\iint_\Omega |u(\tau,s)|^2\,d\tau\,ds$$

and it follows that $\|\mathscr{L}^{-1}\|\leqq M$.

If, on the other hand, $\|\mathscr{L}^{-1}\| \leqq M'$, consider the result of applying $\mathscr{L}^{-1}$ to a separable $u(\cdot,\cdot)$:

$$u(t,s)=u(\tau)\cdot v(s), \qquad t,s \geqq 0,$$
$$=0 \qquad \text{otherwise,}$$

with $u(\cdot)$, $v(\cdot) \in L_2(0,\infty)$. Since

$$\int\int_\Omega \left| \int_\Sigma^t g(t,\tau;s)u(\tau,s)d\tau \right|^2 dt\,ds \leqq M' \int\int |u(\tau,s)|^2\, d\tau\, ds$$

we have for all $u(\cdot)$ of compact support, say supp $u(\cdot)=[0,N]$,

$$\int_0^T \int |v^2|(s) \left| \int_0^N g(t,\tau;s)\,u(\tau)\,d\tau \right|^2 dt\,ds \leqq M' \int_0^\infty |v(s)|^2\,ds \int_0^N |u(\tau)|^2\,d\tau.$$

Now, rather than a fixed function $v(\cdot)$ in the above, take a sequence $\{v_n(\cdot)\}$ such that $|v_n(\cdot)|^2$ approaches as a distribution a δ-function at s_0. Since $g(t,\tau;\cdot)$ is continuous, we conclude that

$$\int_0^T \left| \int_0^N g(t,\tau;s_0)\,u(\tau)\,d\tau \right|^2 dt \leqq M' \int_0^N |u(\tau)|^2\,d\tau,$$

and hence that

$$\int_0^\infty \left| \int_0^\infty g(t,\tau;s_0)\,u(\tau)\,d\tau \right|^2 dt \leqq M' \|u(\cdot)\|^2.$$

This shows L_{s_0} has a bounded inverse, $\|L_{s_0}^{-1}\| \leqq (M')^{1/2}$, and by the corollary to Lemma 1 it follows that L^{-1} exists.

With Lemma 2 above, we have established the equivalence of the invertibility of our original operator L with almost-periodic coefficients to the invertibility of the partial differential operator $\mathscr{L}$ with multiply-periodic coefficients. While the transform method of Lemma 5 below may be applied directly to the operator $\mathscr{L}$, the results so obtained do not appear computationally useful. For this reason, we further modify $\mathscr{L}$ as follows.

Let $r(\cdot)$ be a polynomial of degree m with the property that all the roots of $r(\cdot)$ have strictly negative real parts. We define the partial differential operator $\mathscr{RL}$ by the conditions that the domain of $\mathscr{RL}$ is $\mathscr{D}(\mathscr{RL})=\{x(\cdot,\cdot)\in L_2(\Omega)$ such that the distributional derivatives $(\partial^{i+j}/(\partial t^i \partial s^j))x(\cdot,\cdot)\in L_2(\Omega)$, for $0 \leqq i \leqq n$, $0 \leqq j \leqq m$, and $x|_\Sigma = (\partial x/\partial t)|_\Sigma = \cdots = (\partial^{n-1}x/\partial t^{n-1})|_\Sigma = 0$, $\mathscr{L}x|_\Sigma = (\partial/\partial s)\mathscr{L}x|_\Sigma = \cdots = (\partial^{m-1}\mathscr{L}x/\partial s^{m-1})|_\Sigma = 0\}$, while $\mathscr{RL}$ is defined on its domain by

$$\mathscr{RL}\,x(t,s)=r\left(\frac{\partial}{\partial s}\right)\left[p\left(\frac{\partial}{\partial t}\right)+\{k_1(\omega_1 t+\omega_2 s)+k_2(\omega_2 t+\omega_1 s)\}q\left(\frac{\partial}{\partial t}\right)\right]x(t,s).$$

Under the above definition $\mathscr{RL}$ is a closed operator, and if $\Gamma(s,\sigma;t)$ denotes the Green's function associated with the differential equation

$$r\left(\frac{\partial}{\partial s}\right)x(t,s)=u(t,s),$$

subject to the boundary conditions $x|_\Sigma = \cdots = (\partial^{m-1}x/\partial s^{(m-1)})|_\Sigma = 0$, then the fundamental solution associated with the operator $\mathscr{RL}$ is just the composition of $g(\cdot)$ with $\Gamma(\cdot)$.

LEMMA 3. *Let L be defined as in Lemma* 1. *Then L has a bounded inverse on* $L_2(0, \infty)$ *if and only if* $\mathcal{RL}$ *has a bounded inverse on* $L_2(\Omega)$.

Proof. If L has a bounded inverse, then $\mathcal{L}^{-1}\mathcal{R}^{-1}$ is an inverse for $\mathcal{RL}$, and $\mathcal{R}^{-1}$ is bounded by the assumption on the roots of $r(\cdot)$.

Conversely, if $\mathcal{RL}$ has a bounded inverse, then again choosing $u(t, s) = u(t)\, v(s)$ as in Lemma 2 we obtain the inequality

$$\int_0^\infty \int_0^\infty \left| \int_0^t g(t, \tau; s)\, u(\tau)\, d\tau \right|^2 \cdot \left| \int_0^s \varphi(s-\sigma)\, v(\sigma)\, d\sigma \right|^2 ds\, dt$$

$$\leqq M \int_0^\infty |u(\tau)|^2\, d\tau \cdot \int_0^\infty |v(\sigma)|^2\, d\sigma.$$

The fact that the Green's function associated with $\mathcal{R}$ appears with argument $(s-\sigma)$ is a consequence of the assumption that $v(\cdot)$ vanishes for negative values of its argument.

Choosing a $v(\cdot)$ such that

$$\left| \int_0^s \varphi(s-\sigma)\, v(\sigma)\, d\sigma \right|^2 > 0$$

almost everywhere and applying Fubini's Theorem, we conclude that

$$\int_0^\infty \left| \int_0^t g(t, \tau; s)\, u(\tau)\, d\tau \right|^2 dt < \infty$$

for almost all s. This does not complete the proof, however, as the set of measure zero in question might depend on the choice of $u(\cdot)$.

This difficulty is circumvented by noting that the control system described by the ordinary differential equation (3) together with the "output" $x(t)$ is both uniformly controllable and uniformly observable by [12].

This means that there exist a finite set of controls $\{u_1(\cdot) \cdots u_n(\cdot)\}$ driving the system from the origin to an n-dimensional set of vectors at time 1. Combing this with the above we have

$$\int_0^\infty \left| \int_0^1 g(t, \tau; s)\, u_i(\tau)\, d\tau \right|^2 dt < \infty$$

for almost all s, and all i, $1 \leqq i \leqq n$, and hence

$$\sum_{i=1}^N \int_1^\infty \left| \int_0^1 g(t, \tau; s)\, u_i(\tau) d\tau \right|^2 dt < \infty.$$

By uniform observability, we have

$$\sum_{i=1}^n \int_1^\infty \left| \int_0^1 g(t, \tau; s)\, u_i(\tau)\, d\tau \right|^2 dt \geqq k \int_1^\infty \|\varphi(t, 1; s)\|^2\, dt$$

where $\varphi(t, 1; s)$ denotes the transition matrix associated with the "parametrized" differential equation

$$\left[p\left(\frac{d}{dt}\right) + \{k_1(\omega_1 t + \omega_2 s) + k_2(\omega_2 t + \omega_1 s)\} q\left(\frac{d}{dt}\right) \right] x(t) = 0.$$

Using the relation between the above equation and the "translated" version of (3) as in Lemma 1, we conclude that with $\varphi(t, t_0)$ the transition matrix associated with

(3) we have

$$\int_{t_0}^{\infty} \|\varphi(t, t_0)\|^2 \, dt < M$$

for almost all t_0 first, then for all t_0 by continuity. By [4, p. 190], this implies exponential stability for (3) from which the existence of L^{-1} readily follows.

We next reduce the problem to one of the invertibility of a partial differential operator on a quarter-plane.

The change of variables

$$\tau = \omega_1 t + \omega_2 s$$

$$\sigma = \omega_2 t + \omega_1 s$$

naturally introduces a linear operator

$$T: L_2(\Omega) \to L_2(\pi^+ = \{\tau, \sigma \geqq 0\})$$

with a bounded inverse; also we have the formal relations

$$T \frac{\partial}{\partial t} T^{-1} = \omega_1 \frac{\partial}{\partial \tau} + \omega_2 \frac{\partial}{\partial \sigma}$$

$$T \frac{\partial}{\partial s} T^{-1} = \omega_2 \frac{\partial}{\partial \tau} + \omega_1 \frac{\partial}{\partial \sigma}.$$

We define the operator $R(P+KQ)$ as $T(\mathscr{R}\mathscr{L})T^{-1}$. Obviously we have

$$R(P+KQ)x(\tau, \sigma) = r\left(\omega \frac{\partial}{\partial \tau} + \omega_1 \frac{\partial}{\partial \sigma}\right) \cdot$$

$$\left[p\left(\omega_1 \frac{\partial}{\partial \tau} + \omega_2 \frac{\partial}{\partial \sigma}\right) + \{k_1(\tau) + k_2(\sigma)\} q\left(\omega_1 \frac{\partial}{\partial \tau} + \omega_2 \frac{\partial}{\partial \sigma}\right)\right] x(\tau, \sigma),$$

and the description of the domain of $R(P+KQ)$ follows readily from that of $\mathscr{R}\mathscr{L}$.

LEMMA 4. *Let L be as in Lemma 1, and $R(P+KQ)$ as described above. Then L has a bounded inverse on $L_2(0, \infty)$ if and only if $R(P+KQ)$ has a bounded inverse on $L_2(\pi^+)$.*

We next introduce what is essentially a Fourier–Laplace transform operator in order to reduce the problem of analysis of $R(P+KQ)$ on $L_2(\pi^+)$ to an equivalent problem in the L_2 space of a compact domain.

The basic methods are essentially similar to those in [5].

Given a function $x(\cdot, \cdot) \in L_2(\pi^+)$, we define a transform according to

$$\hat{x}(\tau, \sigma; z_1, z_2) = \sum_{n,m=1}^{\infty} x(\tau+n, \sigma+m) z_1^{-n} z_2^{-m}.$$

This expression may be regarded as the Laplace transform of an $L_2([-1, 0]^2)$-valued function defined on $Z^+ \times Z^+$ by associating to the point (n, m) the restriction of the function $x(\cdot, \cdot)$ to the square $[n-1, n] \times [m-1, m] \subset R^2$.

The problem which remains is to describe the action of the operator $R(P+KQ)$ in terms of the action induced on the transform function $\hat{x}(\cdot, \cdot; z_1, z_2)$. This may most easily be described in terms of the "conventional" Laplace transform on $L_2(\pi^+)$. In fact,

$$\int_{-1}^{0}\int_{-1}^{0} e^{-2\pi ik\tau} e^{-2\pi il\sigma} z_1^{-\tau} z_2^{-\sigma}\hat{x}(\tau,\sigma;z_1,z_2)\,d\tau\,d\sigma = \hat{X}(2\pi ik+\operatorname{Log} z_1, 2\pi il+\operatorname{Log} z_2), \tag{13}$$

with $\hat{X}(\cdot,\cdot)$ the usual Laplace transform of $x(\cdot,\cdot)\in L_2(\pi^+)$. Because of the fact that the operator $R(P+KQ)$ has essentially been constructed as a composition of minimal ordinary differential operators, the action of $R(P+KQ)$ is readily described in terms of the transform functions $\hat{x}$.

If we let H denote the linear mapping H: $x\to\hat{x}$ described above, and further define

$$T(z_1,z_2)=HR(P+KQ)H^{-1}$$
$$(T(z_1,z_2)\colon \mathcal{D}\subset L_2([-1,0]^2)\to L_2([-1,0]^2)),$$

then from the observation that a minimal constant coefficient ordinary differential operator on a half-line is diagonalized by the Laplace transform we conclude that

(i) $\mathcal{D}(T(z_1,z_2))=H\mathcal{D}(R(P+KQ))$,
(ii) $z_1^{-\tau}z_2^{-\sigma}HQx=Q_1(z_1,z_2)z_1^{-\tau}z_2^{-\sigma}Hx$,
where $Q_1(z_1,z_2)$ acting in $L_2([-1,0])$ is diagonalized by the Fourier transform (series):

$$Q_1(z_1,z_2)e^{-2\pi ik\tau}e^{-2\pi il\sigma}=g(\omega_1(2\pi ik+\operatorname{Log} z_1)+\omega_2(2\pi il+\operatorname{Log} z_2))e^{-2\pi ik\tau}e^{-2\pi il\sigma},$$

$$\text{(iii)}\quad \begin{aligned} z_1^{-\tau}z_2^{-\sigma}HPx&=P_1(z_1,z_2)z_1^{-\tau}z_2^{-\sigma}Hx,\\ z_1^{-\tau}z_2^{-\sigma}HRx&=R_1(z_1,z_2)z_1^{-\tau}z_2^{-\sigma}Hx,\end{aligned} \tag{14}$$

with R_1, P_1 defined in a manner analogous to Q_1 above.

From the definition of the operator H it follows that K acts on Hx (and also on $z_1^{-\tau}z_2^{-\sigma}Hx$) as simple multiplications by the function $K(\cdot,\cdot)$: $K(\tau,\sigma)=k_1(\tau)+k_2(\sigma)$.

Because of the relatively simple description of the operators R_1, Q_1, P_1 and K given above, it is convenient below to deal with the operator

$$T_1(z_1,z_2)\colon \mathcal{D}_{T_1}\subset L_2([-1,0]^2)\to L_2([-1,0]^2)$$

defined by

$$T_1(z_1,z_2)=R_1(z_1,z_2)(P_1(z_1,z_2)+KQ_1(z_1,z_2)).$$

LEMMA 5. *For a fixed* (z_1,z_2), $T_1(z_1,z_2)$ *has a bounded inverse if and only if* $T(z_1,z_2)$ *has a bounded inverse.*

Proof. This follows immediately from the preceding definitions.

Remark. For fixed (z_1,z_2), $T_1(z_1,z_2)$ may be identified with the operator arising from the imposition of periodic boundary conditions in connection with the formal partial differential operator

$$r\left(\omega_2\left(\frac{\partial}{\partial\tau}+\operatorname{Log} z_1\right)+\omega_1\left(\frac{\partial}{\partial\sigma}+\operatorname{Log} z_2\right)\right)$$
$$\cdot\left[p\left(\omega_1\left(\frac{\partial}{\partial\tau}+\operatorname{Log} z_1\right)+\omega_2\left(\frac{\partial}{\partial\sigma}+\operatorname{Log} z_2\right)\right)\right.$$
$$\left.+K(\tau,\sigma)q\left(\omega_1\left(\frac{\partial}{\partial\tau}+\operatorname{Log} z_1\right)+\omega_2\left(\frac{\partial}{\partial\sigma}+\operatorname{Log} z_2\right)\right)\right].$$

With the abbreviations $\boldsymbol{\omega}=(\omega_1, \omega_2)$, $\boldsymbol{\omega}'=(\omega_2, \omega_1)$, $\operatorname{Log} \mathbf{z}=(\operatorname{Log} z_1, \operatorname{Log} z_2)$, the above may be written as

$$r(\boldsymbol{\omega}' \cdot (\boldsymbol{\nabla}+\operatorname{Log} \mathbf{z}))[p(\boldsymbol{\omega} \cdot (\boldsymbol{\nabla}+\operatorname{Log} \mathbf{z}))+K(\tau, \sigma)q(\boldsymbol{\omega} \cdot (\boldsymbol{\nabla}+\operatorname{Log} \mathbf{z}))]$$

LEMMA 6. *There exists $R>0$ such that $\forall |z_1|, |z_2|>R$, the operator $T_1(\mathbf{z})$ considered as an operator on $L_2([-1, 0]^2)$ has a bounded inverse.*

Proof. Rearrange $R(P+KQ)(\mathbf{z})$ as

$$r(\boldsymbol{\omega}' \cdot (\boldsymbol{\nabla}+\operatorname{Log} \mathbf{z}))\circ\left[1+K(\tau, \sigma)\frac{q}{p}(\boldsymbol{\omega} \cdot (\boldsymbol{\nabla}+\operatorname{Log} \mathbf{z}))\right]$$

$$\circ p(\boldsymbol{\omega} \cdot (\boldsymbol{\nabla}+\operatorname{Log} \mathbf{z}))$$

and note that

(i) $r(\boldsymbol{\omega}' \cdot (\boldsymbol{\nabla}+\operatorname{Log} \mathbf{z}))$ is invertible $\forall |z_1|, |z_2|>1$,

(ii) $p(\boldsymbol{\omega} \cdot (\boldsymbol{\nabla}+\operatorname{Log} \mathbf{z}))$ is invertible for $\operatorname{Re} \boldsymbol{\omega} \cdot \operatorname{Log} \mathbf{z}$ greater than the real part of any zero of the polynomial $p(\cdot)$, and

(iii) $\|q/p(\boldsymbol{\omega} \cdot (\boldsymbol{\nabla}+\operatorname{Log} \mathbf{z}))\| \to 0$ as $\operatorname{Re} \boldsymbol{\omega} \cdot \operatorname{Log} \mathbf{z} \to \infty$, so that the middle factor is invertible for $|z_1|, |z_2|>R$ sufficiently large.

Invertibility of $R(P+KQ)(\mathbf{z})$ then follows for $|z_1|, |z_2|>R$.

Remark. The intuitive content of Lemma 6 is simply that the Laplace transform of the solution $x(\cdot, \cdot)$ to the equation

$$R(P+KQ)\,x(\tau, \sigma)=u(\tau, \sigma)$$

exists for values of the transform variables having sufficiently large real part.

LEMMA 7. *Let $\mathbf{z}$ be such that $T_1(\mathbf{z})$ has a bounded inverse on $L_2([-1, 0]^2)$. Then T_1^{-1} is a compact operator; in fact $T_1^{-1}(\mathbf{z})$ is of Hilbert–Schmidt type.*

Proof. Using the fact that it suffices to prove that $(T_1(\mathbf{z})-\lambda)^{-1}$ is compact (Hilbert–Schmidt) for one value of λ ([13, p. 210]), and the fact that for large $|\lambda|$ we have $\|RKQ(RP-\lambda)^{-1}\|$ small, we see that it is sufficient to establish the result with $K=0$.

With $K=0$, the operator $R_1P_1(\mathbf{z})$ is diagonalized by the standard orthonormal basis, so that an explicit formula for the Hilbert–Schmidt norm of $(R_1P_1(\mathbf{z}))^{-1}$ may be written simply by summing the squares of the eigenvalues of $(R_1\,P_1\,(\mathbf{z}))^{-1}$. In the case that $\operatorname{Log} \mathbf{z}$ and the zeroes of $r(\cdot)$, $p(\cdot)$ are all real, the explicit formula for (the candidate for) the square of the Hilbert–Schmidt norm is just

$$\sum_{j,k=0}^{\infty} \frac{1}{\prod_{l,m}(A_l^2+4\pi^2(\omega_2 k+\omega_1 j)^2)(B_m^2+4\pi^2(\omega_1 k+\omega_2 j)^2)},$$

where $\{A_l\}$ and $\{B_m\}$ are given by the zeroes of $p(\cdot)$ and $r(\cdot)$ plus $\operatorname{Log} z_1$ and $\operatorname{Log} z_2$ respectively.

The above sum may be readily estimated in terms of a convergent double integral, showing that $T_1(\mathbf{z})^{-1}$ is Hilbert–Schmidt. In the case that either a zero of $p(\cdot)$ or $r(\cdot)$, or the term $\operatorname{Log} \mathbf{z}$ has a nonzero imaginary part the above formula becomes somewhat more complicated; however, no essential complication is introduced.

LEMMA 8. *If $T_1^{-1}(\mathbf{z})$ exists, for some $\mathbf{z}^0$, $|z_1^0|, |z_2^0| \geqq 1$, then $T_1^{-1}(\mathbf{z})$ is analytic in a neighborhood of $\mathbf{z}^0$.*

Proof. Identify $T_1(\mathbf{z})$ with the differential operator

$$r(\boldsymbol{\omega}' \cdot (\boldsymbol{\nabla}+\operatorname{Log} \mathbf{z}))[p(\boldsymbol{\omega} \cdot (\boldsymbol{\nabla}+\operatorname{Log} \mathbf{z}))+k(\tau, \sigma)q(\boldsymbol{\omega} \cdot (\boldsymbol{\nabla}+\operatorname{Log} \mathbf{z}))],$$

subject to periodic boundary conditions, and split the operator as $A+B$, with

$$A \sim r(\boldsymbol{\omega}' \cdot \nabla + 1)p(\boldsymbol{\omega} \cdot \nabla + 1),$$
$$B = T_1(z) - A.$$

Then the "perturbation" B is A-bounded, B is analytic, and $I + BA^{-1}$ is analytic and bounded. It then follows [15, Sect. VII. 6] that $T_1^{-1}(\mathbf{z})$ is analytic.

From Lemmas 6, 7 and 8 we conclude that the multiply-periodic partial differential equation

$$R(P+KQ)\,x(\tau, \sigma) = u(\tau, \sigma), \qquad \tau, \sigma \geqq 0,$$

subject to the boundary conditions noted above may be solved in terms of the "H-transforms" defined above to give (first for $|z_1|$, $|z_2| > R$) (since $R(P+KG)$ is readily seen to be $1-1$ by construction) the unique solution

$$\hat{x}(\tau, \sigma; z_1, z_2) = (R(P+KQ)(\mathbf{z}))^{-1}\hat{u}(\tau, \sigma; z_1, z_2).$$

To ensure that $\hat{x}(\,\cdot\,, \cdot\,; z_1, z_2)$ so determined actually is the "H-transform" of a function in $L_2(\pi^+)$, we must invoke an operator-valued version of the Paley–Wiener Theorem ([11, Thm. 2]). This gives the following result.

THEOREM 1. *Let L be as in Lemma* 1, *and $T_1(\mathbf{z})$ as in Lemma* 8. *Then $T_1^{-1}(\mathbf{z})$ exists as an analytic operator valued function of Hilbert–Schmidt class for all $|z_1|$, $|z_2|$ sufficiently large. Further, the operator L has a bounded inverse on $L_2(0, \infty)$ if and only if $T_1^{-1}(\mathbf{z})$ has an analytic extension over all $|z_1|$, $|z_2| \geqq 1$.*

Proof. The fact that $T_1^{-1}(\mathbf{z})$ is analytic and of Hilbert–Schmidt type wherever it exists was shown above.

Next, the operator $R(P+KQ)$ originally defined on $L_2(\pi^+)$ naturally extends (again because of its definition in terms of the corresponding minimal ordinary differential operators) to an operator on $L_2(R_2)$. Because of the uniqueness of solutions to the equation

$$R(P+KQ)\,x(\tau, \sigma) = u(\tau, \sigma)$$

for functions $u(\,\cdot\,, \cdot\,)$ of compact support, the methods of [3] are readily adapted to the situation under study. Essentially the only modification required is to replace the functions supported on intervals in [3] by functions supported on rectangles.

We then conclude (cf. [3, Thm. 4.1]) that the operator $R(P+KQ)$ has a bounded inverse on $L_2(\pi^+)$ if and only if the "natural extension" of $R(P+KQ)$ to $L_2(R_2)$ described above has a *causal* bounded inverse.

Viewing $L_2(R_2)$ (and $L_2(\pi_+)$) as spaces of $L_2([-1, 0]^2)$-valued functions defined on $Z \times Z$ ($Z^+ \times Z^+$) and using the Laplace transform operator H, the causality condition may be replaced by an equivalent analyticity requirement. It follows from the vector-valued version of the Paley–Weiner Theorem given in [11, Thm. 2] that $R(P+KQ)$ has a causal inverse if and only if $T^{-1}(\mathbf{z})$ has a bounded analytic extension over the region $|z_1|$, $|z_2| \geqq 1$.

To complete the proof, we must show that the analyticity condition may be restated in terms of the operator $T_1(\mathbf{z})$; that is, that $T^{-1}(\mathbf{z})$ has an analytic extension if and only if the same holds true for $T_1^{-1}(\mathbf{z})$.

First, if $T^{-1}(\mathbf{z})$ exists and is analytic, then by Lemma 5 $T_1^{-1}(\mathbf{z})$ exists; by Lemma 7 $T_1^{-1}(\mathbf{z})$ is analytic in a neighborhood of $\mathbf{z}$.

On the other hand, if $T_1^{-1}(\mathbf{z})$ has an analytic extension for $|z_1|$, $|z_2| > 1$, bounded for $|z_1|$, $|z_2| \geqq 1$, then from (12) and the definition of $T_1(\mathbf{z})$ it follows that the conventional) Laplace transform of the function $x(\cdot,\cdot)$

$$x(t,s)=[z_1^\tau z_2^\sigma H^{-1}T_1^{-1}z_1^{-\tau}z_2^{-\sigma}H]u(t,s)$$

is analytic over the dual of the cone $\pi^+\subset R^2$. Further, this transform has square integrable boundary values; by the (scalar-valued) Paley–Weiner theorem $x(\cdot,\cdot)\in L_2(\pi^+)$. Hence $R(P+KQ)$ has a causal bounded inverse, and so $T^{-1}(\mathbf{z})$ must possess a bounded analytic extension.

Theorem 1 essentially reduces the question of invertibility of the operator L to that of the location of the locus of the spectrum of the two-parameter family of operators $T_1(\mathbf{z})$, where for each fixed $\mathbf{z}$, $T_1(\mathbf{z})$ has a compact resolvent.

From a practical or computational point of view it would be useful, if possible, to reduce the dimensionality of this parametrization. The idea that such a reduction might be possible is intuitively due to the fact that the dimensionality of the problem has been "artificially" increased by the introduction of the variable "s" whose only real function is to render the problem computationally more tractable. In order to state conveniently Theorem 2 below, it is useful to introduce yet another partial differential operator. $T_2(s)$ is defined in analogy to $T_1(\mathbf{z})$ as

$$r(\boldsymbol{\omega}'\cdot\nabla)[p(\boldsymbol{\omega}\cdot\nabla+s)+K(\tau,\sigma)\,q(\boldsymbol{\omega}\cdot\nabla+s)]$$

acting in $L_2([-1,0]^2)$ and subject to periodic boundary conditions.

THEOREM 2. *Let L be as in Lemma* 1, *and the family of operators* $T_2(s)$ *as defined above. Then L has a bounded inverse on* $L_2(0,\infty)$ *if and only if* $0\notin\sigma(T_2(s))$, $\operatorname{Re}(s)\geqq 0$.

Proof. By Theorem 1 and Lemma 8, L has a bounded inverse if and only if $T_1(\mathbf{z})$ has an inverse for all $|z_1|, |z_2|\geqq 1$. By Lemma 6, we see that in fact the invertibility of $T_1(\mathbf{z})$ need only be verified for the compact set $1\leqq|z_1|, |z_2|\leqq R$, for some $R>1$.

Consider the operators

$$\begin{aligned} M(\mathbf{z}) &= r(\boldsymbol{\omega}'\cdot\nabla)[r(\boldsymbol{\omega}'\cdot(\nabla+\operatorname{Log}\mathbf{z}))]^{-1}\\ &= [r(\boldsymbol{\omega}'\cdot(\nabla+\operatorname{Log}\mathbf{z}))]^{-1}r(\boldsymbol{\omega}'\cdot\nabla)\\ &= [R_1(\mathbf{z})]^{-1}[R_1(1,1)]. \end{aligned}$$

By the standing assumption on $r(\cdot)$, $M(\mathbf{z})$ exists for $|z_1|, |z_2|\geqq 1$; if $\mathbf{z}$ is restricted to $1\leqq|z_1|, |z_2|\leqq R$, we have also the result that $M^{-1}(\mathbf{z})$ is uniformly bounded:

$$\|M^{-1}(\mathbf{z})\|\leqq N<\infty.$$

It then follows that $T_1(\cdot)$ is invertible at the point $\mathbf{z}$ if and only if $M(\mathbf{z})\,T_1(\cdot)$ is invertible at $\mathbf{z}$. This last operator may be interpreted as the closure of

$$r(\boldsymbol{\omega}'\cdot\nabla)[p(\boldsymbol{\omega}\cdot(\nabla+\operatorname{Log}\mathbf{z}))+K(\tau,\sigma)\,q(\boldsymbol{\omega}\cdot(\nabla+\operatorname{Log}\mathbf{z}))] \tag{15}$$

subject to periodic boundary conditions; to complete the proof it remains only to justify the replacement of the term $\boldsymbol{\omega}\cdot\operatorname{Log}\mathbf{z}$ with the single variable s.

Consider the collection of unitary operators U_{kj}; $L_2([-1,0]^2)\to L_2([-1,0]^2)$:

$$U_{kj}\varphi(\tau,\sigma)=e^{-2\pi ik\tau}e^{-2\pi ij\sigma}\varphi(\tau,\sigma).$$

Introduction of the operator U_{kj} as a similarity has the effect of replacing the operator (15) by (the closure of)

$$\begin{aligned} r(\boldsymbol{\omega}'\cdot\nabla+2\pi ik\omega_2+2\pi ij\omega_1)\cdot[p+k(\tau,\sigma)q]&\\ \cdot(\boldsymbol{\omega}\cdot\nabla+\omega_1 2\pi ij+\omega_2 2\pi ik+\boldsymbol{\omega}\cdot\operatorname{Log}\mathbf{z})& \end{aligned}$$

(subject to periodic boundary conditions).

Since ω_1 and ω_2 are irrationally related, $\{\omega_1 j+\omega_2 k\}$ is dense in R_1, and so

$$\bigcup_{k,j,|z|\geqq 1} s=(2\pi i(\omega_1 k+\omega_2 j)+\boldsymbol{\omega}\cdot \operatorname{Log} \mathbf{z})$$

is just the half-plane Re $(s)\geqq 0$.

Since the presence of the "imaginary shift in the argument of $r(\cdot)$ has no effect on the invertibility of the product (essentially the same argument as above), the conclusion follows.

The versions of these results given above in § 1 differ from these essentially in that complex variables s_1 and s_2 of nonnegative real part replace the Log z_1 and Log z_2 appearing above, while $L_2([0,1]^2)$ replaces $L_2([-1,0]^2)$ as the space under consideration. The possibility of eliminating the logarithms in the statement of Theorem 1 follows from the density argument employed in the proof of Theorem 2. The replacement of $L_2([-1,0]^2)$ by $L_2([0,1]^2)$ follows from the observation that, with doubly periodic coefficients and boundary conditions, the problem is essentially posed on the 2-torus (or more simply by the change of variables $\tau'=\tau+1$, $\sigma'=\sigma+1$).

4. An example of the results. In this section we apply the results obtained above to a typical example.

Consider the system described by the integral equation

$$e(t)+\{\sin(2\pi t)+\sin(2\pi\sqrt{2}t)\} \cdot \int_0^t e^{-(1/2)(t-\tau)} \sin(t-\tau)\, e(\tau)\, d\tau = u(t), \qquad t\geqq 0.$$

This represents a second order system with an almost-periodic feedback gain given by $k(t)=\sin(2\pi t)+\sin(2\pi\sqrt{2}t)$.

The equivalent differential equation representation of the above system is just

$$[D^2+D+\tfrac{5}{4}+\{\sin(2\pi t)+\sin(2\pi\sqrt{2}t)\}]x(t)=u(t), \qquad t\geqq 0$$

together with the initial conditions $x(0)=Dx(0)=0$. This identifies the polynomials and parameters of the problem as

$$p(D)=D^2+D+\tfrac{5}{4},$$
$$q(D)=1,$$
$$\omega_1=1,$$
$$\omega_2=\sqrt{2}.$$

The polynomial $r(D)$ is arbitrary (subject to the stability constraint), and may be taken as

$$r(D)=D+1.$$

The explicit form of the differential expression corresponding to the operator $T_2(s)$ of Theorem 2 is

$$r(\boldsymbol{\omega}'\cdot\nabla)[p(\boldsymbol{\omega}\cdot\nabla+s)+K(\tau,\sigma)q(\boldsymbol{\omega}\cdot\nabla+s)]\,x(\tau,\sigma)$$
$$=\left[\sqrt{2}\frac{\partial}{\partial\tau}+\frac{\partial}{\partial\sigma}+1\right] \cdot\left[\left(\frac{\partial}{\partial\tau}+\sqrt{2}\frac{\partial}{\partial\sigma}+s\right)^2+\left(\frac{\partial}{\partial\tau}+\sqrt{2}\frac{\partial}{\partial\sigma}+s\right)+\tfrac{5}{4} +\{\sin 2\pi\tau+\sin 2\pi\sigma\}\right]x(\tau,\sigma).$$

According to Theorem 2, stability of the original system is determined by the location of the spectrum of $T_2(s)$ as s ranges over $\mathrm{Re}\,(s) \geqq 0$.

We remark that the proof of Lemma 6 shows that only a subset of the region $\mathrm{Re}\,(s) \geqq 0$ need be checked in this regard, since $T_2(s)$ is invertible for $\mathrm{Re}\,(s)$ sufficiently large. The location of this critical region of the s-plane may be determined by an application of a slight extension of the so-called circle criterion (see for example, [4, pp. 218–227]). Applying this to the example under consideration, we deduce that $T_2(s)$ is invertible for all s such that

$$\sup_{\omega \in R} \left| \frac{2}{(i\omega + s)^2 + (i\omega + s) + \frac{5}{4}} \right| < 1.$$

The above expresion may be easily manipulated to yield the "safe" region for s. Similar considerations show that in connection with the use of Theorem 1, only z_1, z_2 such that

$$\sup_{\omega \in R} \left| \frac{2}{(i\omega + \mathrm{Log}\, z_1 + \sqrt{2}\, \mathrm{Log}\, z_2) + (i\omega + \mathrm{Log}\, z_1 + \sqrt{2}\, \mathrm{Log}\, z_2) + \frac{5}{4}} \right| < 1$$

need be considered. This reduces consideration to a compact region of C^2, which is a computational necessity for practical use of these results.

5. Conclusions. The results of Theorem 1 and 2 show that it is indeed possible to obtain necessary and sufficient conditions for invertibility of certain almost-periodic differential operators in terms of the locus of the spectra of a parametrized family of related operators having a compact resolvent.

The methods employed are conceptually simple, being based on the two ideas of multiple time scales and vector-valued Laplace transforms. The complications which emerge when these simple ideas are applied, on the other hand, seem to us somewhat messy. We suspect strongly that this may be an inherent aspect of the problem.

The work reported above raises several questions, among which we mention the following.

In attempting to approximate the resolvent $T_2^{-1}(s)$ by finite dimensional operators the questions of rates of convergence and error estimates are relevant. Since the polynomial $r(\cdot)$ is largely arbitrary, is it possible to exploit this to accelerate convergence?

The above derivations have been carried out in terms of two independent frequencies, but the corresponding results of any finite number of frequencies follow readily from a modification of the methods given above. It is clear, however, that the amount of computation required to make approximate stability calculations increases geometrically with the number of frequencies. Is it possible to choose the arbitrary (analog of) $r(\cdot)$ in such a way the numerical burden is reduced?

This work was undertaken partly in the hope that a Fredholm index theory somehow analogous to the results in [5] would emerge in the process. We suspect that the condition that $T_2(s)$ be invertible for all $\mathrm{Re}\, s = 0$ is a form of the condition that L be Fredholm, although we have made no attempt to verify this. Constant coefficient examples show that the index in that case is not simply related to the Fredholm determinant of $T_2^{-1}(i\omega)$, which might be regarded as a candidate for a function providing a winding number.

Acknowledgment. The author would like to thank the referees for their careful and useful comments on an earlier version of this paper.

REFERENCES

[1] G. ZAMES, *On the input-output stability of non-linear time-varying feedback systems, Part* I, IEEE Trans Automatic Control, AC-11 (1966), pp. 228–238.

[2] J. C. WILLEMS, *On the asymptotic stability of the null solution of linear differential equations with periodic coefficients*, Ibid., AC-13 (1968), pp. 65–72.

[3] ———, *Stability, instability, invertibility, and causality*, SIAM Journal on Control, 7 (1969), pp. 645–671.

[4] R. W. BROCKETT, *Finite Dimensional Linear Systems*, John Wiley, New York, 1970.

[5] J. H. DAVIS, *Fredholm operators, encirclements and stability criteria*, SIAM Journal on Control, 10 (1972), pp. 608–628.

[6] J. C. WILLEMS, *The Analysis of Feedback Systems*, M.I.T. Press, Cambridge, MA, 1971.

[7] S. GOLDBERG, *Unbounded Linear Operators*, McGraw-Hill, New York, 1966.

[8] M. VAN DYKE, *Perturbation Methods in Fluid Mechanics*, Academic Press, New York, 1964.

[9] J. L. MASSERA AND J. J. SCHAEFFER, *Linear Differential Equations and Function Spaces*, Academic Press, New York, 1966.

[10] R. J. SACHER AND G. R. SELL, *Existence of dichotomies and invariant splittings for linear differential systems* I, Journal Differential Equations, 15 (1974).

[11] Y. FOURES AND I. E. SEGAL, *Causality and analylicity*, Transactions Amer. Math. Soc. 78 (1965), pp. 385–405.

[12] L. M. SILVERMAN AND B. D. O. ANDERSON, *Controllability, observability and stability of linear systems*, SIAM Journal on Control, 6 (1968), pp. 121–129.

[13] E. HILLE AND R. S. PHILLIPS, *Functional Analysis and Semi-Groups*, American Mathematical Society Colloquium publications XXXI, Providence, RI, 1957.

[14] N. DUNFORD AND J. T. SCHWARTZ, *Linear Operators*, vol. 1, Wiley Interscience, New York, 1967.

[15] D. R. SMITH, *The multivariable method in singular perturbation analysis*, SIAM Review, 17 (1975), pp. 221–273.

SIAM. J. MATH. ANAL.
Vol. 9, No. 6, December 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0002 $01.00/0

ANALYTIC REPRESENTATIONS AND FOURIER TRANSFORMS OF ANALYTIC FUNCTIONALS IN Z' CARRIED BY THE REAL SPACE*

J. W. DE ROEVER†

Abstract. In the space Z', the Fourier transform of the space $\mathscr{D}'$ of Schwartz-distributions, the notion of carrier is introduced. A characterization is given of all distributions in $\mathscr{D}'$, the Fourier transform of which is carried by $\mathbb{R}^n$. Both, such distributions and the analytic functionals in Z' carried by $\mathbb{R}^n$, are represented as sum of boundary values of holomorphic functions. This extends the case of tempered distributions which, regarded as elements of Z', are obviously carried by $\mathbb{R}^n$.

1. Introduction. In Vladimirov [11, 26.3 and 26.4, Thm. 2] theorems are derived concerning Fourier transforms of tempered distributions in S' with support in a certain, unbounded, convex set. These Fourier transforms can be represented as boundary values in S' of holomorphic functions in a tubular, radial domain. This yields an analytic representation of distributions in S', i.e., a sum of distributional boundary values of certain holomorphic functions. In this paper these notions are generalized such that the boundary values are no longer attained in S' but in the space Z', which is the Fourier transform of the space $\mathscr{D}'$ of Schwartz-distributions. These boundary values are analytic functionals in Z' carried by $\mathbb{R}^n$ and their inverse Fourier transforms are distributions in $\mathscr{D}'$. Moreover, a characterization is obtained of those distributions in $\mathscr{D}'$ such that their Fourier transforms are all analytic functionals carried by $\mathbb{R}^n$. Then analytic representations of such functionals and distributions are given. This extends the case of tempered distributions which, regarded as elements of Z', are obviously carried by $\mathbb{R}^n$. Finally, the spaces of functions, holomorphic in tubular radial domains of exponential type in Im z and of polynomial growth in Re z, and the spaces of their inverse Fourier transforms are provided with topologies such that Fourier transformation is a topological isomorphism.

2. Notations and definitions. We will denote vectors in $\mathbb{C}^n$ by $z=(z_1,\cdots,z_n)=x+iy$ and by $\zeta=\xi+i\eta$, where x, y, ξ, η are vectors in $\mathbb{R}^n$. The norm in $\mathbb{R}^n$ and $\mathbb{C}^n$ will be denoted as $\|\cdot\|$. For $t, w\in\mathbb{R}^n$ or $\mathbb{C}^n$, $t\cdot w$ will stand for $t_1w_1+\cdots+t_nw_n$. Let α be an n-tuple nonnegative integers; then $D_x^\alpha=D_1^{\alpha_1}\cdots D_n^{\alpha_n}$, where $D_j=\partial/\partial x_j$, and $\vec{D}_x=(\partial/\partial x_1,\cdots,\partial/\partial x_n)$; when no confusion arises the subscript x will be omitted. Similarly, for $t\in\mathbb{R}^n$ or $\mathbb{C}^n$ t^α is defined. Furthermore, $|\alpha|=\alpha_1+\cdots+\alpha_n$ and $\alpha!=\alpha_1!\cdots\alpha_n!$. For $b\in\mathbb{R}$, $\bar{b}$ will denote the vector $(b,\cdots,b)\in\mathbb{R}^n$.

We recall the testfunction spaces $\mathscr{E}$, $\mathscr{D}$ (cf. [9], $\mathscr{K}$ in [4]), S and Z and their duals denoted by $'$, which also refers to the strong topology. The action of an element $f\in W'$ on functions $\phi\in W$ will be denoted by $\langle f,\phi\rangle$ or $\langle f,\phi\rangle_{W'}$. Sometimes we will write f_ξ and $\phi(\xi)$, if W consists of functions of ξ. In that case W', W and $\langle\cdot,\cdot\rangle$ will be denoted by W'_ξ, W_ξ and $\langle\cdot,\cdot\rangle_\xi$, too. We mention explicitly the action of a function f, regarded as an element of $\mathscr{D}'$, to a testfunction ϕ in $\mathscr{D}$:

$$\langle f,\phi\rangle=\int_{\mathbb{R}^n} f(\xi)\phi(\xi)\,d\xi. \tag{2.1}$$

The Fourier transform of an L^1-function ϕ is given by

$$\mathscr{F}[\phi](x)=\mathscr{F}[\phi(\xi)](x)=\int_{\mathbb{R}^n}\phi(\xi)\,e^{ix\cdot\xi}\,d\xi.$$

* Received by the editors November 26, 1975, and in revised form February 9, 1977.
† Department of Applied Mathematics, Mathematisch Centrum, Amsterdam, the Netherlands.

Then $\mathscr{D}$ is the Fourier transform of Z and $\mathscr{D}(a)$ of $Z(a)$, where $\mathscr{D}(a)$ is the space of C^∞-functions with support in $\{\xi \mid \|\xi\| \leqq a\}$. $Z(a)$ consists of entire functions ψ, such that for all m

$$(2.2) \qquad \|\psi\|_m \overset{\text{def}}{=} \sup_{z \in \mathbb{C}^n} (1+\|z\|)^m e^{-a\|y\|} |\psi(z)| < \infty.$$

The Fourier transform of a distribution $f \in \mathscr{D}'$ is defined as that element $\mathscr{F}[f]$ of Z' for which

$$(2.3) \qquad \langle \mathscr{F}[f], \psi \rangle_{Z'} = \langle f, \mathscr{F}[\psi] \rangle_{\mathscr{D}'}, \qquad \psi \in Z;$$

cf. Schwartz [9].[1]

Elements μ of $Z'(a)$ can be written as $\langle \mu, \psi \rangle = \int f(x)\psi(x)\,dx$ for some entire function f; see Gel'fand and Shilov [4, III, § 2.3]. Hence μ is a functional on the space of restrictions to $\mathbb{R}^n$ of functions in $Z(a)$. In general, this is no longer true when $\mu \in Z'$. We say that $\mu \in Z'$ is *carried by the closed set* $\Omega \subset \mathbb{C}^n$, if for every ε-neighborhood $\Omega(\varepsilon)$ of Ω μ is already a functional on the space $Z|_{\Omega(\varepsilon)}$ of restrictions to $\Omega(\varepsilon)$ of functions in Z, where $Z|_{\Omega(\varepsilon)}$ carries the topology induced by Z, i.e., in (2.2) the supremum should be taken over all $z \in \Omega(\varepsilon)$. According to Ehrenpreis [2, Thm. 5.13*] a fundamental system of neighborhoods of zero in Z is given by $V(k, \alpha) = \{\psi \in Z \mid |\psi(z)| \leqq \alpha k(z)\}$, where $\alpha > 0$ and k is a positive continuous function of the following form: let $\{a_j\}$ be a strictly increasing sequence of integers with $a_0 = a_1 = a_2 = 0$, $a_{j+2} > 2a_j$, and let l be a positive integer; set $k(z) = (1+\|x\|)^{-l}(1+\|y\|)^{-l} \exp((j-2)\|y\|)$ for $a_j(1+\log(1+\|x\|)) \leqq \|y\| \leqq \frac{1}{2}a_{j+1}(1+\log(1+\|x\|))$; the definition of k is completed by requiring that k is a function of $\|x\|, \|y\|$ which is continuous and such that, for fixed $\|x\|, \log k(\|x\|, \|y\|) + l[\log(1+\|x\|) + \log(1+\|y\|)]$ is linear in $\|y\|$ in the regions in which it is not already defined above. Then a fundamental system of neighborhoods of zero in $Z|_{\Omega(\varepsilon)}$ is obtained by $\{\psi \in Z|_{\Omega(\varepsilon)} \mid |\psi(z)| \leqq \alpha k(z), z \in \Omega(\varepsilon)\}$. Now the Hahn–Banach theorem and Riesz' representation theorem imply that for every $\varepsilon > 0$ an analytic functional μ carried by Ω can be represented as a measure μ_ε on $\Omega(\varepsilon)$ satisfying

$$\int_{\Omega(\varepsilon)} k_\varepsilon(z) |d\mu_\varepsilon(z)| \leqq K_\varepsilon$$

where k_ε is a function as described above depending on ε.

In this paper we will be concerned with closed sets Ω which are bounded in the imaginary direction; i.e., Ω is contained in a set of the form $\{z \mid \|y\| \leqq b\}$, $b \leqq 0$. Then, if $Z(a)_m$ denotes the normed space of functions in $Z(a)$ provided with the norm (2.2), the spaces

$$Z_F = \operatorname*{proj\,lim}_{m\to\infty} \operatorname*{ind\,lim}_{a\to\infty} Z(a)_m \quad \text{and} \quad Z = \operatorname*{ind\,lim}_{a\to\infty} \operatorname*{proj\,lim}_{m\to\infty} Z(a)_m$$

induce the same topology on $Z|_{\Omega(\varepsilon)}$. Indeed, according to Ehrenpreis [2, Thm. 5.10] a fundamental system of neighborhoods of zero in Z_F is given by $V(k', \alpha)$, where now $k'(z) = (1+\|x\|)^{-m} k_1'(y)$ with $m \geqq 0$ and with k_1' a positive, continuous, function

[1] Since also Z is the Fourier transform of $\mathscr{D}$, a similar definition gives the Fourier transform of analytic functionals $\mu \in Z'$. Then, if for $\mu \in Z'$ $\check{\mu}$ is defined by $\langle \check{\mu}_\zeta, \psi(\zeta) \rangle = (2\pi)^{-n} \langle \mu_\xi, \overline{\psi(\bar{\zeta})} \rangle$, $\psi \in Z$, $g = \mathscr{F}[\mu]$ implies $\mathscr{F}[g] = \check{\mu}$. Therefore, the theorems of § 4 of this paper dealing with inverse Fourier transforms g of elements $\mu \in Z'$ may just as well have been formulated with $g = \mathscr{F}[\mu]$, hence with Fourier transforms of analytic functionals instead of inverse Fourier transforms; cf. the title of this paper.

dominating every exp $a\|y\|$, $a>0$. Z_F is the Fourier transform of the space $\mathscr{D}_F$, the test space for the finite order distributions. Hence the (inverse) Fourier transforms of all elements μ in Z' carried by Ω are finite order distributions and, moreover, for every $\varepsilon>0$ these μ satisfy

$$|\langle \mu, \psi\rangle| \leqq K_\varepsilon \sup_{z\in\Omega(\varepsilon)} (1+\|x\|)^{m(\varepsilon)}|\psi(z)|, \qquad \psi \in Z,$$

with K_ε and $m(\varepsilon)$ depending on ε and μ. The above representation yields that for every $\varepsilon>0$ μ can be represented as a measure μ_ε on $\Omega(\varepsilon)$ satisfying

$$\int_{\Omega(\varepsilon)} \frac{|d\mu_\varepsilon(z)|}{(1+\|x\|)^{m(\varepsilon)}} \leqq K_\varepsilon. \tag{2.4}$$

In particular, we will be concerned with analytic functionals in Z' carried by $\mathbb{R}^n$.

The support of a distribution $f\in W'$, where W is a space of C^∞-functions on $\mathbb{R}^n$, is defined as the smallest closed set U such that any $\xi_0 \notin U$ has an open neighborhood V_0 with $\langle f, \phi\rangle = 0$ for every $\phi\in W$ with $\phi(\xi)=0$ if $\xi\notin V_0$. In § 5 we will show that for certain sets U, in particular for convex sets, f can be represented as sum of weak derivatives of measures on U.

Finally, C will denote an open cone in $\mathbb{R}^n$ (i.e., $t\in C \Rightarrow \lambda t\in C, \lambda>0$), ch($C$) its convex hull, pr $C=\{y\in C \mid \|y\|=1\}$ and $C'\Subset C$ means that C' is a relatively compact subcone of C, i.e., $\overline{\text{pr } C'}\subset \text{pr } C$. The function

$$\mu_C(\xi) = \sup_{y\in \text{pr } C} -y\cdot\xi$$

is the indicatrix of C. C^* denotes the closed cone $\{\xi \mid y\cdot\xi \geqq 0, y\in C\} = \{\xi \mid \mu_C(\xi)\leqq 0\}$ and $C_* = \mathbb{R}^n\backslash C^* = \{\xi \mid \mu_C(\xi)>0\}$. Furthermore

$$\rho_C = \sup_{\xi\in\mathbb{R}^n} (\mu_{\text{ch}(C)}(\xi)/\mu_C(\xi)).$$

We will consider holomorphic functions f in the tubular radial domain $T^C=\mathbb{R}^n+iC$ if C is connected, or in the tubular radial set T^C if C is not connected. In the former case we say that $f(z)$ has a boundary value f^* in Z' or S' as $y\to 0$, $y\in C$ or $y \in C'\Subset C$, respectively, if for all $\phi\in Z$ or S the limit

$$\langle f^*, \phi\rangle = \lim_{\substack{y\to 0\\ y\in C \text{ or } C'}} \int_{\mathbb{R}^n} f(x+iy)\phi(x)\,dx \tag{2.5}$$

exists. The boundary value in S' is said to be attained on the distinguished boundary of T^C, i.e., on the set $\{z \mid z\in T^C, \text{Im } z=0\}$. However, if the limit exists in Z', it is less clear a "boundary value on the distinguished boundary," as we will see in § 3.

3. Boundary values in Z' and S'. This section is concerned with results obtained in de Roever [6, § 8]. Here, these results will be given in more detail and their consequences will be examined more closely.

We consider functions f holomorphic in the tube domain $T^B=\mathbb{R}^n+iB$, where B is a domain in $\mathbb{R}^n$. For any $y\in B$ $f(z)=f(x+iy)$ is a C^∞-function in x, i.e., $f(z)\in\mathscr{E}_x$. We regard $\mathscr{E}$ as the strong dual of $\mathscr{E}'$, the space of distributions with compact support. By the Paley–Wiener–Schwartz theorem the Fourier transform of $\mathscr{E}'$ is known as the space H of entire functions of exponential type in Im ζ and of polynomial growth in Re ζ, provided with the topology such that the Fourier transformation is a topological isomorphism between $\mathscr{E}'$ and H. Then $\mathscr{E}$ is the Fourier transform of the dual H' of H. With these definitions we have the following lemma (cf. de Roever [6, Lemma 8.1]).

LEMMA 3.1. *Let f be a holomorphic function in T^B with B a domain in $\mathbb{R}^n$ and let $y_0 \in B$. Then for all y with $y + y_0 \in B$*

$$f(x+iy+iy_0) = \mathscr{F}[e^{-y\cdot\zeta}\mu(y_0)_\zeta](x)$$

where

$$\mu(y_0) = \mathscr{F}^{-1}[f(x+iy_0)] \in H'.$$

Proof. Since this lemma is shown already in [6, Lemma 8.1] we will not give all the details.

With the aid of Cauchy's formula it can be shown that for $\|y\|$ sufficiently small

$$\sum_{k=0}^{\infty} \frac{(iy\cdot \vec{D}_x)^k}{k!} f(x+iy_0) = f(x+iy+iy_0)$$

converges in $\mathscr{E}$. For $\phi_x \in \mathscr{E}'$ let $\psi(\zeta) = \mathscr{F}[\phi_x](\zeta) \in H$. Then we have

$$\begin{aligned}\langle \phi_x, f(x+iy+iy_0)\rangle_{\mathscr{E}'} &= \sum_{k=0}^{\infty} \left\langle \mu(y_0)_\zeta, \frac{(-y\cdot\zeta)^k}{k!}\psi(\zeta)\right\rangle_{H'} \\ &= \lim_{N\to\infty} \left\langle \sum_{k=0}^{N} \frac{(-y\cdot\zeta)^k}{k!}\mu(y_0)_\zeta, \psi(\zeta)\right\rangle_{H'}.\end{aligned}$$

Hence the weak limit for $N \to \infty$ exists in H'. Since H is a Montel space the strong limit exists and equals

$$\sum_{k=0}^{\infty} \frac{(-y\cdot\zeta)^k}{k!}\mu(y_0)_\zeta = e^{-y\cdot\zeta}\mu(y_0)_\zeta.$$

Thus

$$f(x+iy+iy_0) = \mathscr{F}[e^{-y\cdot\zeta}\mu(y_0)_\zeta](x)$$

for y with $\|y\|$ sufficiently small. By analytic continuation this formula holds for every y with $y + y_0 \in B$. □

Since $\mathscr{D}$ is dense in $\mathscr{E}'$, Z is dense in H. Hence for $f \in \mathscr{E}$ the Fourier transform $\mathscr{F}[f]$ is also determined by (2.3). The space of restrictions to $\mathbb{R}^n$ of functions in Z is dense in S, so that if, moreover, f belongs to S' then (2.3) implies

$$\langle \mathscr{F}[f], \psi\rangle_{S'} = \langle f, \mathscr{F}[\psi]\rangle_{S'}, \qquad \psi \in S.$$

Thus we have obtained the following corollary.

COROLLARY 3.2. *If $f(z)$, holomorphic in T^B, for each $y \in B$ belongs to S'_x, then with $y_0 \in B$ and with $g(y_0) = \mathscr{F}^{-1}[f(x+iy_0)] \in S'$*

$$e^{-y\cdot\xi}g(y_0)_\xi \in S'_\xi$$

for y such that $y + y_0 \in B$ and

$$f(z+iy_0) = \mathscr{F}[e^{-y\cdot\xi}g(y_0)_\xi](x).$$

From this corollary one derives as in Vladimirov [11, Thm. 26.1] that $f(z)$ belongs to S'_x for each $y \in B$ if and only if it satisfies

$$|f(z)| \leqq M(K)(1+\|x\|)^{m(K)}, \qquad y \in K \Subset B \tag{3.1}$$

for all compact sets K in B, where $M(K)$ and $m(K)$ depend on these sets K. Moreover, Corollary 3.2 and Vladimirov [11, 26.2] yield that, if $f(z)$ belongs to S'_x for each $y \in B$, it necessarily satisfies, for every α,

$$|D^\alpha f(z)| \leqq M'(K)(1+\|x\|)^{m'(K)}, \qquad y \in K \Subset \mathrm{ch}(B), \tag{3.2}$$

for all compact sets K in the convex hull $\mathrm{ch}(B)$ of B, where $M'(K)$ and $m'(K)$ depend on K and α. Hence $f(z)$ belongs to S'_x for each $y \in B$ means that (2.1) holds for $\phi \in S$. Clearly, in that case for each $y \in B$ $f(z)$ belongs to Z' under definition (2.1), too.

Next we consider a function $f(z)$, holomorphic in T^B, which for each $y \in B$ as a function of x belongs to Z'. This means that $f(z)$ is a continuous linear functional on the space of restrictions to $\mathbb{R}^n$ of functions in Z, where this space carries the topology of S. Its closure is S, hence $f(z)$ belongs to S'_x for each $y \in B$. Thus $f(x+iy) \in S'_x$ is equivalent to $f(x+iy) \in Z'$ and this should be interpreted in the sense of definition (2.1) for $\phi \in S$ or Z.

Let us now consider the limit in Z' as y tends to zero. From (3.1) it follows that for any $y \in B$ and y_0 such that $y + y_0 \in B$

$$\langle f(z), \psi(x)\rangle = \int_{\mathbb{R}^n} f(z+iy_0)\psi(x+iy_0)\,dx, \qquad \psi \in Z.$$

Therefore, the limit f^* as $y \to 0$ of $f(x+iy)$ exists in Z' and it is given by

$$\langle f^*, \psi\rangle \overset{\text{def}}{=} \lim_{y\to 0} \int_{\mathbb{R}^n} f(x+iy+iy_0)\psi(x+iy_0)\,dx = \int_{\mathbb{R}^n} f(x+iy_0)\psi(x+iy_0)\,dx \tag{3.3}$$

for all $\psi \in Z$ and $y_0 \in B$. This limit is independent of $y_0 \in B$. Note, that when $0 \notin \bar{B}$ f^* can never exist in S'. If $0 \in \bar{B}$, $0 \notin B$, we may call f^* the boundary value in Z' according to (2.5) and this boundary value is independent of the path $y \to 0$ in B. Still f^* might not belong to S' nor satisfy definition (2.1). For example, take $B = \{y \mid y > 0\}$ in $\mathbb{R}^1$ and $f(z) = \exp 1/z$. This function satisfies (3.1), but $\int [\exp(1/x)]\psi(x)\,dx$, $\psi \in Z$, does not exist. In general, it follows from (3.1) and (3.3) that f^* is an element of Z' carried by $\mathbb{R}^n$ (see § 2 for the definition of carrier in Z').

The inverse Fourier transform $g = \mathcal{F}^{-1}[f^*]$ is an element of $\mathcal{D}'$. For $\phi \in \mathcal{D}$ and $y_0 \in B$ from (3.3) we derive

$$\begin{aligned}\langle g, \phi\rangle_{\mathcal{D}'} &= \langle f^*, \psi\rangle_{Z'} = \langle f(x+iy_0), \psi(x+iy_0)\rangle_{Z'}\\ &= \langle g(y_0)_\xi, e^{y_0\cdot\xi}\phi(\xi)\rangle_{\mathcal{D}'} = \langle e^{y_0\cdot\xi} g(y_0)_\xi, \phi(\xi)\rangle_{\mathcal{D}'},\end{aligned}$$

where $\psi = \mathcal{F}^{-1}[\phi]$ and $g(y_0) = \mathcal{F}^{-1}[f(x+iy_0)]$. Hence

$$g_\xi = e^{y\cdot\xi} g(y)_\xi, \qquad y \in B, \tag{3.4}$$

is independent of y and we have obtained (as in Corollary 3.2)

$$e^{-y\cdot\xi} g_\xi \in S'_\xi, \qquad y \in B, \tag{3.5}$$

and

$$f(z) = \mathcal{F}[e^{-y\cdot\xi} g_\xi](x). \tag{3.6}$$

Conversely, for a distribution $g \in \mathcal{D}'$ satisfying (3.5) the function (3.6) satisfies (3.1), hence also (3.2) (see Vladimirov [11, Thm. 26.2]). Conditions (3.5) and (3.6) then also hold for $y \in \mathrm{ch}(B)$.

The distribution $g \in \mathcal{D}'$ can be obtained as follows: for $y \in K$

$$(3.7)\quad g_\xi = e^{y\cdot\xi}\mathcal{F}^{-1}[f(x+iy)]_\xi = (A(K)-\vec{D}\cdot\vec{D})^{m'(K)}\frac{1}{(2\pi)^n}\int \frac{f(x+iy)\,e^{-iz\cdot\xi}}{(A(K)+z\cdot z)^{m'(K)}}\,dx,$$

where $m'(K) \geqq \frac{1}{2}(m(K)+n+1)$ and where $A(K)$ is so large that for $y \in K$

$$|A(K)+z\cdot z| \geqq 1 + x\cdot x.$$

The integral is independent of y. Hence for every $K \Subset B$ there are constants $m(K), M_\alpha(K)$ and continuous functions $g_{\alpha,K}$ on $\mathbb{R}^n$, $|\alpha| \leqq m(K)$, such that g can be represented as

$$(3.8)\quad \begin{aligned} g_\xi &= \sum_{|\alpha|\leqq m(K)} D^\alpha g_{\alpha,K}(\xi),\\ |g_{\alpha,K}(\xi)| &\leqq M_\alpha(K)\,e^{y\cdot\xi},\quad \forall y \in K.\end{aligned}$$

In that case g also satisfies (3.8) for every $K \Subset \mathrm{ch}(B)$. Conversely, if a distribution $g \in \mathcal{D}'$ satisfies (3.8) then (3.5) holds.

Next we consider the case that the limit f^* exists in S'. Let C be an open connected cone in $\mathbb{R}^n$ and let for each $C' \Subset C$ $R(C')$ be a positive number depending on C'. Let B be a domain in $\mathbb{R}^n$ containing each set $\{y \mid y \in \overline{C'}, 0<\|y\|\leqq R(C')\}$. Let f be a holomorphic function in T^B which satisfies a stronger condition than (3.1), namely

$$(3.9)\quad |f(z)| \leqq M(C')(1+\|x\|)^{m(C')}(1+\|y\|^{-k}),\qquad y \in C',\quad \|y\| \leqq R(C')$$

for each compact subcone C' of C and for some k and $m(C')$ depending on C'. We may let k depend on C', too, but in Lemma 3.3 it will be shown that (3.9) is satisfied for a fixed k anyhow. Then the limit

$$f^* = \lim_{\substack{y\to 0\\ y\in C'}} f(x+iy)$$

exists in S'_x and it is independent of C' and the path $y \to 0$ in C'; see Vladimirov [11, Thm. 26.3]. Here the most general case arises if $R(C')$ tends to zero as C' approaches C. Now f^* is called the boundary value of $f(z)$ in S' on the distinguished boundary. Clearly f^* is attained in Z', too, but if the limit exists in Z' only, (3.3) shows that the boundary value in Z' may be concentrated on other sets than the distinguished boundary as well.

Let $g = \mathcal{F}^{-1}[f^*]$, where f^* is the boundary value in S' of a function $f(z)$ satisfying (3.9); then g belongs to S'. The representation (3.7) now holds for $y \in \overline{C'}$, $0<\|y\|\leqq R(C')$ and for $m'(K) \geqq \frac{1}{2}(m(C')+n+1)$ and $A(K) \geqq R(C')^2+1$. Also here the integral is independent of y. Therefore, for any $\xi \in \mathbb{R}^n$ we can choose a suitable $y = y_\xi$. For $\xi \in C'^*$ and $\|\xi\| \geqq 1/R(C')$ we choose $y_\xi \in C'$ with $\|y_\xi\| = 1/\|\xi\|$. Then $y_\xi \cdot \xi \leqq 1$ and $0 \leqq y\cdot\xi$ for all $y \in C'$, $\|y\| \leqq R(C')$. For $\xi \in C'_*$ we have $\min_{y\in \mathrm{pr}\, C'} y\cdot\xi < 0$. Let the minimum be attained for y'_ξ; then we take $y_\xi = R(C')y'_\xi$ and we have

$$y_\xi\cdot\xi \leqq y\cdot\xi,\quad \forall y \in C',\quad \|y\| \leqq R(C').$$

Now we take $y = y_\xi$ in the integral in (3.7) and we find that for every C' there are a positive integer $m(C')$, constants $M_\alpha(C')$ and continuous functions $g_{\alpha,C'}$ on $\mathbb{R}^n$, $|\alpha| \leqq$

$m(C')$, such that g can be represented as

$$g_\xi = \sum_{|\alpha| \leqq m(C')} D^\alpha g_{\alpha,C'}(\xi),$$

(3.10)

$$|g_{\alpha,C'}(\xi)| \leqq M_\alpha(C')(1+\|\xi\|)^k e^{y\cdot\xi}, \quad \forall y \in C', \quad \|y\| \leqq R(C').$$

If $R(C') \leqq R$ for all $C' \Subset C$ and if in (3.9) m is independent of C', then in (3.10) $m(C')$ and the functions $g_{\alpha,C'}$ can be chosen independent of C'.

The following lemma shows that $f(z)$ satisfying (3.9) for $C' \Subset C$ satisfies (3.9) also for $C' \Subset \mathrm{ch}(C)$, hence that $g \in S'$ satisfying (3.10) satisfies (3.10) also for $C' \Subset \mathrm{ch}(C)$.

LEMMA 3.3. *Let C be an open connected cone in $\mathbb{R}^n$, let for each compact subcone C' of C $R(C')$ be a positive number and let B be a domain in $\mathbb{R}^n$ containing every set $\{y \mid y \in \overline{C'}, 0 < \|y\| \leqq R(C')\}$. If $f(z)$ is a holomorphic function in T^B that satisfies* (3.1) *such that the limit f^* in Z' as $y \to 0$ belongs to S', then $f(z)$ attains this limit already in S' as $y \to 0$, $y \in C' \Subset \mathrm{ch}(C)$ and, moreover, f satisfies* (3.9) *for each $C' \Subset \mathrm{ch}(C)$ and $R(C')$ such that*

(3.11) $$\{y \mid y \in \overline{C'}, 0 < \|y\| \leqq R(C')\} \subset \mathrm{ch}(B).$$

If in (3.1) *$m(K)$ does not depend on K, then $m(C')$ in* (3.9) *does not depend on C'.*

Proof. Fix $C' \Subset \mathrm{ch}(C)$ and choose C'' and C''' with $C' \Subset C'' \Subset C''' \Subset \mathrm{ch}(C)$. Fix $R(C''')$ such that (3.11) holds for $C' = C'''$. Let U_0 be an open ε-neighborhood of C''^*; then there is a $\delta > 0$ such that for $y \in C'$, $-y \cdot \xi \leqq -\delta\|y\|\,\|\xi\|$ if $\xi \in U_0$ outside a compact set. Finally choose finitely many vectors $y_j \in \mathrm{pr}\, C'''$ and positive numbers $\delta_j < 1$ such that the open sets

$$U_j = \{\xi \mid y_j \cdot \xi < -\delta_j\|\xi\|\} \cap \{\xi \mid -y \cdot \xi < 2\delta_j\|\xi\|,\ y \in \mathrm{pr}\, C'\},$$

$j = 1, \cdots, p$, cover C''_*. Let $\{\lambda_j\}_{j=0}^p$ be a partition of unity subordinate to the covering $\bigcup_{j=0}^p U_j$ of $\mathbb{R}^n$, such that λ_j is a multiplier in S' for every j. Now for all $y \in C'$, $\|y\| \leqq \frac{1}{4}R(C''')$ the functions

$$\lambda_j(\xi)\, e^{R(C''')y_j\cdot\xi - y\cdot\xi}, \qquad j = 0, \cdots, p, \quad y_0 = 0,$$

belong to S. In Lemma 5.2 it will be shown that the Fourier transforms of $\lambda_j(\xi)\, e^{-y\cdot\xi} g_\xi$, $j = 0, \cdots, p$, where $g \in S'$ satisfies (3.5), are equal to

$$\langle e^{-R(C''')y_j\cdot\xi} g_\xi, \lambda_j(\xi)\, e^{R(C''')y_j\cdot\xi + iz\cdot\xi}\rangle_{S'}, \qquad j = 0, 1, \cdots, p,$$

respectively. Hence if $g = \mathcal{F}^{-1}[f^*]$, we obtain for $y \in C'$, $\|y\| \leqq \frac{1}{4}R(C''')$

$$\begin{aligned}|f(z)| &= |\mathcal{F}[e^{-y\cdot\xi} g_\xi]| = \left|\mathcal{F}\left[\sum_{j=0}^p \lambda_j(\xi)\, e^{-y\cdot\xi} g_\xi\right]\right| \\ &\leqq |\langle g_\xi, \lambda_0(\xi)\, e^{iz\cdot\xi}\rangle| + \sum_{j=1}^p |\langle e^{-R(C''')y_j\cdot\xi} g_\xi, \lambda_j(\xi)\, e^{R(C''')y_j\cdot\xi + iz\cdot\xi}\rangle| \\ &\leqq M_1(1+\|z\|)^m \sup_{t \geqq 0} (1+t)^k e^{-\delta\|y\|t} + M_2(C''')(1+\|z\|)^{m'(C''')} \\ &\leqq M(C')(1+\|x\|)^{m(C')}(1+\|y\|^{-k})\end{aligned}$$

for some M_1, m and k depending on g and some $M_2(C''')$ depending on $M(K)$ in (3.1) and $m'(C''')$ depending on $m(K)$ in (3.1), where $K = \{y \mid y \in C''', \|y\| = R(C''')\}$. Together with (3.1) this yields (3.9) for $C' \Subset \mathrm{ch}(C)$.

Furthermore, it now follows that g can be represented as in (3.10), hence that the set $\{e^{-y\cdot\xi}g_\xi | y \in C', \|y\| \leqq R(C')\}$ is bounded in S', where $R(C')$ is such that (3.11) holds. Since the limit as $y \to 0$ of $e^{-y\cdot\xi}g_\xi$ exists in $\mathscr{D}'$, this limit exists in S' as $y \to 0$, $y \in C'$. Hence $f(x+iy) \to f^*$ in S' as $y \to 0$, $y \in C' \Subset \mathrm{ch}(C)$. □

Altogether we have obtained the following theorem.

THEOREM 3.4. *Let f be a holomorphic function in T^B satisfying* (3.1), *where B is a domain in $\mathbb{R}^n$. Then it satisfies* (3.2), *too. The limit f^* of $f(z)$ as $y \to 0$ exists in Z' and its inverse Fourier transform g satisfies* (3.8) *for every $K \Subset \mathrm{ch}(B)$. Conversely, a distribution $g \in \mathscr{D}'$ satisfying* (3.8) *for $K \Subset B$ satisfies* (3.8) *for $K \Subset \mathrm{ch}(B)$, too, and then the function* (3.6), *which is defined because* (3.5) *holds, satisfies* (3.2). *Moreover, if* (3.11) *holds, a function f holomorphic in T^B has a boundary value f^* in S', provided that f satisfies* (3.9), *which then is satisfied for $C' \Subset \mathrm{ch}(C)$, too. Then the inverse Fourier transform g of f^* satisfies* (3.10) *for every $C' \Subset \mathrm{ch}(C)$. Conversely, a distribution $g \in S'$ satisfying* (3.10) *for $C' \Subset C$ satisfies* (3.10) *for $C' \Subset \mathrm{ch}(C)$, too, and then the function* (3.6), *which is defined because* (3.5) *holds, satisfies* (3.9) *for $C' \Subset \mathrm{ch}(C)$. There is no mixture of these cases, i.e., a holomorphic function f in T^B, B such that* (3.11) *holds, satisfying* (3.1) *and having a boundary value f^* in Z', which is an element of S', already satisfies* (3.9).

We conclude this section with an example of a function f that satisfies (3.1). Let $B = \{y | y > 0\} \subset \mathbb{R}^1$ and let

$$f(z) = z^{(i/\cos z + \cos z)}.$$

For some positive constants A and B

$$\begin{aligned} |f(z)| &\leqq M(y)(1+|x|)^{A/\sinh y + B\cosh y} \\ &\leqq M(r,R)(1+|x|)^{A/r + B\cosh R}, \qquad 0 < r \leqq y \leqq R. \end{aligned}$$

Let $m(r, R)$ be an integer larger than $A/r + B \cosh R + 2$; then for all $r \leqq y \leqq R$

$$\begin{aligned} g_\xi &= e^{y\cdot\xi}\mathscr{F}^{-1}[f(x+iy)]_\xi = (iD_\xi)^{m(r,R)}\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{z^{i/\cos z + \cos z}\, e^{-iz\xi}}{z^{m(r,R)}}\,dx \\ &= (D_\xi)^{m(r,R)} g_{r,R}(\xi) \in \mathscr{D}'. \end{aligned}$$

The continuous function $g_{r,R}(\xi)$ satisfies, for all $r \leqq y \leqq R$,

$$|g_{r,R}(\xi)| \leqq M'(r,R)\, e^{y\cdot\xi},$$

i.e.,

$$\begin{aligned} |g_{r,R}(\xi)| &\leqq M'(r,R)\, e^{r\xi}, \qquad \xi > 0, \\ |g_{r,R}(\xi)| &\leqq M'(r,R)\, e^{-R|\xi|}, \qquad \xi < 0. \end{aligned}$$

4. Analytic representations of real carried analytic functionals and of their Fourier transforms. In § 3 we have discussed the behavior of a function $f(z)$ for $\|y\|$ small and we have seen that it does not change if we deal with a domain T^B or its convex hull $T^{\mathrm{ch}(B)}$. In this section we will discuss it for $\|y\|$ large, namely we will consider holomorphic functions $f(z)$ in T^C of exponential type in $\|y\|$. Then $f(z)$ will not satisfy the same estimates for $y \in C$ or for $y \in \mathrm{ch}(C)$, but they differ by a factor ρ_C in the exponent. The support of the inverse Fourier transform g of the boundary value f^* of such functions f is contained in a certain, convex, set. We will obtain representations of analytic functionals μ in Z', which are carried by $\mathbb{R}^n$, as boundary values of holomorphic functions, and also characterizations and analytic representations of the inverse Fourier transforms $g \in \mathscr{D}'$ of μ. In particular this yields the case that μ and

g both are tempered distributions, which case is treated in Vladimirov [11, 26.4, Thm. 2].

Let f be a holomorphic function in T^C, where C is an open connected cone in $\mathbb{R}^n$. It is shown in Vladimirov [11, 26.6], that, if f satisfies for every $\sigma > 0$

$$|f(z)| \leqq P(C', \sigma)(1+\|z\|)^m(1+\|y\|^{-k}) \exp[(b+\sigma)\|y\|], \qquad y \in C'$$

for some m, k and $b \geqq 0$, then the distribution $g \in S'$ in (3.6) has its support in the set

$$\{\xi | -y \cdot \xi \leqq b, \, y \in \operatorname{pr} C\} = \{\xi | \mu_C(\xi) \leqq b\}. \tag{4.1}$$

The proof does not depend on the behavior of f for $\|y\|$ small, nor on the fact that m is independent of C'; cf. (3.9). Hence the same proof as in Vladimirov [11, 26.6] combined with a proof as in obtaining (3.6) shows that the following theorem is true.

THEOREM 4.1. *Let C be an open and connected cone and let the holomorphic function f in T^C satisfy, for every $r > 0$, $\sigma > 0$ and $C' \Subset C$,*

$$|f(z)| \leqq P(C', r, \sigma)(1+\|z\|)^{m(C',r)} \exp[(b+\sigma)\|y\|], \qquad y \in C', \quad \|y\| \geqq r, \tag{4.2}$$

for some $m(C', r)$ depending on C' and r, some $b \geqq 0$ and some constant $P(C', r, \sigma)$ depending on C', r and σ. Then $f(z) = \mathscr{F}[e^{-y\cdot\xi}g_\xi](x)$ and $\lim_{y\to 0} f(x+iy) = \mathscr{F}[g]$ in Z' for some $g \in \mathscr{D}'$ with support in the set (4.1).

The converse of this theorem follows from Theorem 3.4 and the proof of [11, 26.4, Thm. 2]. We consider distributions g in $\mathscr{D}'$, which are represented as sum of weak derivatives of measures satisfying a condition like (3.8). That any distribution in $\mathscr{D}'$, satisfying (3.5) with $B = C$ and with support in a convex set, can be represented in this way, will be shown in the next section. The advantage of writing g as sum of derivatives of measures is that it enables us to let the σ vanish in (4.2), if C is convex.

THEOREM 4.2. *Let C be an open and connected cone and let g be a distribution in $\mathscr{D}'$, such that for each $C' \Subset C$ and $r > 0$ g can be represented as*

$$g_\xi = \sum_{|\alpha| \leqq m(C',r)} D^\alpha \mu_{\alpha,C',r}(\xi),$$

where the measures $\mu_{\alpha,C',r}$ depending on α, C' and r have their support in the set (4.1) *and satisfy*

$$\int_{\mathbb{R}^n} e^{-y\cdot\xi} |d\mu_{\alpha,C',r}(\xi)| \leqq M_\alpha(C', r), \quad \forall y \in C', \quad \|y\| \geqq r, \tag{4.3}$$

for some positive integers $m(C', r)$ depending on C' and r and for some positive constants $M_\alpha(C', r)$ depending on α, C' and r. Then for $y \in \operatorname{ch}(C)$ the function $f(z) = \mathscr{F}[e^{-y\cdot\xi}g_\xi](x)$ is holomorphic in $T^{\operatorname{ch}(C)}$ and satisfies, for each $C' \Subset \operatorname{ch}(C)$ and $r > 0$,

$$|f(z)| \leqq P(C', r)(1+\|z\|)^{N(C',r)} \exp(\rho_C b\|y\|), \qquad y \in C', \quad \|y\| \geqq r, \tag{4.4}$$

for some positive integers $N(C', r)$ and constants $P(C', r)$ both depending on C' and r. Moreover, $\lim_{y\to 0} f(x+iy) = \mathscr{F}[g]$ in Z' and if $m(C', r)$ is independent of C' and r, then $N(C', r)$, too.

Since the distribution g in Theorem 4.1 satisfies the conditions of Theorem 4.2 (cf. § 3 and the next section) and since $\rho_C = 1$ if C is convex, we obtain the following corollary.

COROLLARY 4.3. *Let the cone C be convex; then a function f that satisfies* (4.2) *satisfies* (4.4) *with $\rho_C = 1$; i.e., in* (4.2) *P is actually independent of σ.*

We now give a characterization of all distributions $g \in \mathscr{D}'$ whose Fourier transforms $\mu \in Z'$ are carried by $\mathbb{R}^n$ (see § 2). We have already seen that if $\mu \in Z'$ is the boundary value of a holomorphic function, it is carried by $\mathbb{R}^n$. Hence it remains to characterize those $g \in \mathcal{D}'$ whose Fourier transforms admit an analytic representation and to show that for every $\mu \in Z'$ which is carried by $\mathbb{R}^n$ $\mathcal{F}^{-1}[\mu]$ satisfies this characterization. Moreover, this yields an analytic representation of elements in Z' carried by $\mathbb{R}^n$.

If a closed, convex, cone C^* in $\mathbb{R}^n$ is the dual of an open cone C in $\mathbb{R}^n$, it does not contain a straight line. We divide $\mathbb{R}^n$ into such cones so that the following properties hold:

$$\mathbb{R}^n = \bigcup_{j=1}^{p} C_j^*, \tag{4.5}$$

$$\operatorname{int} C_j^* \cap \operatorname{int} C_k^* = \emptyset, \qquad j \neq k, \tag{4.6}$$

where each C_j^* is a closed, convex, cone not containing a straight line, while the union of any two cones C_j^* contains a straight line. The last property restricts the number p of cones used, $n+1 \leqq p \leqq 2^n$, and furthermore it states that in some sense the cones C_j^* are as large as possible. Let C_j^* be the dual of the open convex cone C_j; then $C = \bigcup_{j=1}^{p} C_j$ is an open cone. Such a cone C can also be obtained directly as follows: let r open half spaces V_k ($V_k = \{y \mid \xi_k \cdot y > 0\}$ for some $\xi_k \in \mathbb{R}^n$, $\|\xi_k\| = 1$) be given such that $\mathbb{R}^n \backslash \{0\} = \bigcup_{k=1}^{r} V_k$, while $\mathbb{R}^n \backslash \{0\}$ is not covered by the union of any $r-1$ half spaces V_k; hence $n+1 \leqq r \leqq 2n$. Then each C_j is the intersection of n half spaces V_k, i.e.,

$$C = \bigcup \{V_{k_0} \cap \cdots \cap V_{k_{n-1}}\},$$

where the union is taken over all n-tuples $\{k_0, \cdots, k_{n-1}\}$ taken from $\{1, 2, \cdots, r\}$. For example, if $n = 2$ we may take

$$C = \{y \mid 0 < \phi < \tfrac{1}{3}\pi\} \cup \{y \mid \tfrac{2}{3}\pi < \phi < \pi\} \cup \{y \mid \tfrac{4}{3}\pi < \phi < \tfrac{5}{3}\pi\},$$

where $\phi = \tan(y_2/y_1)$, or

$$C = \{y \mid 0 < \phi < \tfrac{1}{2}\pi\} \cup \{y \mid \tfrac{1}{2}\pi < \phi < \pi\} \cup \{y \mid \pi < \phi < \tfrac{3}{2}\pi\} \cup \{y \mid \tfrac{3}{2}\pi < \phi < 2\pi\}.$$

If we write $C' \Subset C$, we mean $C' = \bigcup_{j=1}^{p} C_j'$ with $C_j' \Subset C_j$. Furthermore, let λ_j be the characteristic function of the set C_j^*.

Let $g \in \mathcal{D}'$ be such that for any $C' \Subset C$ and any $\varepsilon > 0$ g can be represented as

$$g_\xi = \sum_{|\alpha| \leqq m(C', \varepsilon)} D^\alpha g_{\alpha, C', \varepsilon}(\xi)$$

where the continuous functions $g_{\alpha, C', \varepsilon}$ satisfy

$$|\lambda_j(\xi) g_{\alpha, C', \varepsilon}(\xi)| \leqq M_{\alpha, j}(C', \varepsilon)\, e^{y \cdot \xi}, \quad \forall y \in C_j', \quad \|y\| \geqq \varepsilon; \tag{4.7}$$

cf. (4.3). Assume that g can be written as

$$g = \sum_{j=1}^{p} g_j, \tag{4.8}$$

where g_j has its support in C_j^* and can be represented as in Theorem 4.2. Then it follows from this theorem that $\mathcal{F}[g]$ is the sum of boundary values in Z' of functions f_j, holomorphic in T^{C_j}, satisfying for all $C_j' \Subset C_j$ and $\varepsilon > 0$

$$|f_j(z)| \leqq P(C_j', \varepsilon)(1 + \|z\|)^{N(C_j', \varepsilon)}, \qquad y \in C_j', \quad \|y\| \geqq \varepsilon,$$

for $j = 1, \cdots, p$, respectively. Hence $\mathcal{F}[g]$ is an element of Z' carried by $\mathbb{R}^n$.

Since for any $C' \Subset C$ there is a positive number $\delta(C') < 1$ such that for $y \in C_j'$ and $\xi \in C_j^*$, $j = 1, \cdots, p$,

$$\delta(C')\|y\|\,\|\xi\| \leqq y \cdot \xi \leqq \|y\|\,\|\xi\|, \tag{4.9}$$

condition (4.7) is equivalent to: for every $\varepsilon > 0$ g can be written as

$$g_\xi = \sum_{|\alpha| \leqq m(\varepsilon)} D^\alpha g_{\alpha,\varepsilon}(\xi),$$

where the continuous functions $g_{\alpha,\varepsilon}$ on $\mathbb{R}^n$ depend on ε and satisfy

$$|g_{\alpha,\varepsilon}(\xi)| \leqq M_\alpha(\varepsilon)\, e^{\varepsilon\|\xi\|}. \tag{4.10}$$

In order to derive (4.8) we introduce the following spaces. Let V be a closed set in $\mathbb{R}^n$, let $E_\varepsilon(V)$ be the Fréchet space of functions ϕ which are C^∞ functions in the interior of V and all of whose derivatives are continuous on V such that the following norms are finite:

$$\|\phi\|_m \sup_{\substack{\xi \in V \\ |\alpha| \leqq m}} e^{\varepsilon\|\xi\|} |D^\alpha \phi(\xi)|, \qquad m = 0, 1, 2, \cdots,$$

and furthermore, let

$$E(V) = \operatorname{ind}\lim_{\varepsilon \to 0} E_\varepsilon(V).$$

Since by (4.5) it is obvious that the continuous restriction map I from $E_\varepsilon(\mathbb{R}^n)$ into $\prod_{j=1}^p E_\varepsilon(C_j^*)$ is injective and has closed image, I is an open mapping according to the open mapping theorem. Hence, it follows from the representation of an open set in an inductive limit space (Floret and Wloka [3, § 23, 3.14]) that also the restriction map from $E(\mathbb{R}^n)$ into $\prod_{j=1}^p E(C_j^*)$ is injective, continuous and open.

In virtue of Treves [10, Prop. 35.4 and Lemma 37.7] the transposed map between the dual spaces is surjective. In particular, the above given $g \in \mathcal{D}'$ actually belongs to $E(\mathbb{R}^n)'$; cf. § 5 and Floret and Wloka [3, § 26, 1.2 and 1.6], and therefore, (4.8) holds with $g_j \in E(C_j^*)'$. This means, exactly, that g_j can be represented as in Theorem 4.2 (cf. § 5).

Next consider an element μ of Z' carried by $\mathbb{R}^n$. As in § 2, for any $\varepsilon > 0$ μ can be represented as a measure μ_ε on an ε-neighborhood $\Omega(\varepsilon)$ of $\mathbb{R}^n$ in $\mathbb{C}^n$ satisfying (2.4). Let μ be the Fourier transform of a distribution $g \in \mathcal{D}'$. Then for any positive $\varepsilon < 1$ and $\phi \in \mathcal{D}$

$$\begin{aligned}\langle g, \phi\rangle &= \int_{\Omega(\varepsilon)} \frac{1}{(2\pi)^n} \int e^{-iz\cdot\xi} \phi(\xi)\, d\xi\, d\mu_\varepsilon(z) \\ &= \int_{\Omega(\varepsilon)} \frac{1}{(2\pi)^n} \int \phi(\xi)(1 - \vec{D}_\xi \cdot \vec{D}_\xi)^{m(\varepsilon)} e^{-iz\cdot\xi}\, d\xi \frac{d\mu_\varepsilon(z)}{(1 + z\cdot z)^{m(\varepsilon)}} \\ &= \int \{(1 - \vec{D}_\xi \cdot \vec{D}_\xi)^{m(\varepsilon)} \phi(\xi)\} \int_{\Omega(\varepsilon)} \frac{e^{-iz\cdot\xi}}{(2\pi)^n (1 + z\cdot z)^{m(\varepsilon)}}\, d\mu_\varepsilon(z)\, d\xi,\end{aligned}$$

hence $g_\xi = (1 - \vec{D}\cdot\vec{D})^{m(\varepsilon)} g_\varepsilon(\xi)$, where g_ε is a continuous function on $\mathbb{R}^n$ which satisfies

$$|g_\varepsilon(\xi)| \leqq \frac{1}{(2\pi)^n} \frac{1}{(1 - \varepsilon^2)^{m(\varepsilon)}} K_\varepsilon\, e^{\varepsilon\|\xi\|},$$

i.e., g satisfies (4.10).

In the following theorem not only $\mathcal{F}[g]$ but also g is represented as sum of boundary values of holomorphic functions.

THEOREM 4.4. *The following four conditions for a distribution* $g \in \mathcal{D}'$ *are equivalent*:

1) $\mathcal{F}[g] \in Z'$ *is carried by* $\mathbb{R}^n$.
2) *For any* $\varepsilon > 0$ g *can be represented as* $g_\xi = \sum_{|\alpha| \leqq m(\varepsilon)} D^\alpha g_{\alpha,\varepsilon}(\xi)$, *where* $g_{\alpha,\varepsilon}$ *are continuous functions on* $\mathbb{R}^n$ *satisfying*
$$|g_{\alpha,\varepsilon}(\xi)| \leqq M_\alpha(\varepsilon)\, e^{\varepsilon\|\xi\|}.$$
3) $\mathcal{F}[g]$ *is the sum of boundary values in* Z' *of functions* f_j *holomorphic in* T^{C_j} *satisfying for any* $C_j' \Subset C_j$ *and any* $\varepsilon > 0$
$$|f_j(z)| \leqq P(C_j', \varepsilon)(1+\|z\|)^{N(C_j', \varepsilon)}, \qquad y \in C_j', \quad \|y\| \geqq \varepsilon$$
for $j = 1, \cdots, p$, *respectively, where* C_j, $j = 1, \cdots, p$, *are any cones satisfying* (4.5) *and* (4.6) *and where* $P(C_j', \varepsilon)$ *and* $N(C_j', \varepsilon)$ *are constants depending on* C_j' *and* ε.
4) g *is the sum of boundary values in* $\mathcal{D}'$ *of holomorphic functions* h_k *in* $T^{\tilde{C}_k}$ *satisfying for any* $\tilde{C}_k' \Subset \tilde{C}_k$ *and any* $\varepsilon > 0$
$$|h_k(\zeta)| \leqq M(\tilde{C}_k', \varepsilon)(1+\|\eta\|^{-m(\varepsilon)})\, e^{\varepsilon\|\xi\|}, \qquad \eta \in \tilde{C}_k'$$
for $k = 1, \cdots, \tilde{p}$, *respectively, where* $\tilde{C}_k$, $k = 1, \cdots, \tilde{p}$, *are any cones satisfying* (4.5) *and* (4.6), *where* $M(\tilde{C}_k', \varepsilon)$ *depends on* $\tilde{C}_k'$ *and* ε, *and where* $m(\varepsilon)$ *depends on* ε *only*.

Proof. We only have to prove that 4) is equivalent to one of the equivalent conditions 1), 2) or 3). First, assume that 4) holds. For $k = 1, \cdots, \tilde{p}$ choose fixed vectors $\eta_k \in \operatorname{pr} \tilde{C}_k'$ for some $\tilde{C}_k' \Subset \tilde{C}_k$. As in Vladimirov [11, proof of Thm. 26.3] for any $\varepsilon > 0$ the absolute values of the functions
$$\int_\tau^1 \int_{\tau_\alpha}^1 \cdots \int_{\tau_1}^1 h_k(\xi + i\tau_0\eta_k)\, d\tau_0\, d\tau_1 \cdots d\tau_\alpha, \qquad k = 1, \cdots, \tilde{p},$$
are bounded by
$$M(\tilde{C}_k', \varepsilon)\, e^{\varepsilon\|\xi\|} \int_\tau^1 \int_{\tau_\alpha}^1 \cdots \int_{\tau_1}^1 (1+\tau_0^{-m(\varepsilon)})\, d\tau_0\, d\tau_1 \cdots d\tau_\alpha$$
and if $\alpha = m(\varepsilon)$, this can be estimated by a constant, depending on ε only, times $\exp(\varepsilon\|\xi\|)$, uniformly for $0 < \tau \leqq 1$. Then integrating another time with respect to τ we get the functions
$$h_{k,\varepsilon}(\xi, \tau) \stackrel{\text{def}}{=} \int_\tau \int_{\tau_{m(\varepsilon)+1}} \cdots \int_{\tau_1} h_k(\xi + i\tau_0\eta_k)\, d\tau_0\, d\tau_1 \cdots d\tau_{m(\varepsilon)+1}, \qquad k = 1, \cdots, \tilde{p},$$
which converge to continuous functions $H_{k,\varepsilon}(\xi, 0)$ of ξ as $\tau \downarrow 0$. The absolute values of these functions are bounded by a constant, depending on ε, times $\exp(\varepsilon\|\xi\|)$. Since
$$\frac{d}{d\tau} h_k(\xi + i\tau\eta_k) = i\eta_k \cdot \vec{D}_\xi h_k(\xi + i\tau\eta_k),$$
we get in distributional sense
$$g_\xi = \sum_{k=1}^{\tilde{p}} \lim_{\tau \downarrow 0} h_k(\xi + i\tau\eta_k) = \sum_{k=1}^{\tilde{p}} \Big\{ (-i\eta_k \cdot \vec{D}_\xi)^{m(\varepsilon)+2} H_{k,\varepsilon}(\xi, 0) + \sum_{j=0}^{m(\varepsilon)+1} \frac{(-i\eta_k \cdot \vec{D}_\xi)^j}{j!} h_k(\xi + i\eta_k) \Big\}.$$
Hence 2) is satisfied.

Now let $g \in \mathscr{D}'$ satisfy 1) and let $\tilde{C} = \bigcup_{k=1}^{\tilde{p}} \tilde{C}_k$ be a cone satisfying (4.5) and (4.6). Since we consider g as a distribution in $\mathscr{D}'$, $\mu = \mathscr{F}[g]$ acts, in principle, on functions in Z. We need a decomposition of μ as sum of analytic functionals carried by $-\tilde{C}_k^*$, just like the decomposition (4.8) of g. For that purpose we introduce the following spaces of analytic functions. If Ω is a closed set in $\mathbb{R}^n$, let Ω_ε be the closed neighborhood in $\mathbb{C}^n$ given by $\Omega_\varepsilon = \{z = x + iy \mid x \in \Omega, \|y\| \leqq \varepsilon\}$; then define the space $Z|_{\Omega\varepsilon}$ by

$$Z\Big|_{\Omega_\varepsilon} = \operatorname*{proj\,lim}_{m\to\infty} Z_m\Big|_{\Omega_\varepsilon},$$

where $Z_m|_{\Omega_\varepsilon}$ is the completion of the set of functions ψ belonging to Z for the norm

$$\sup_{z \in \Omega_\varepsilon} (1 + \|x\|)^m |\psi(z)|,$$

and the space $Z(\Omega)$ by

$$Z(\Omega) = \operatorname*{ind\,lim}_{\varepsilon \to 0} Z|_{\Omega_\varepsilon}.$$

Here, it is not important whether $Z|_{\Omega_\varepsilon}$ consists of all rapidly decreasing functions which are holomorphic in the interior of Ω_ε and continuous in Ω_ε, but we only need that, for $\zeta \in \mathbb{R}^n + i\tilde{C}_k$ and $\Omega = -\tilde{C}_k^*$, $e^{-iz\cdot\zeta}$ as a function of z belongs to $Z|_{\Omega_\varepsilon}$ for every $\varepsilon > 0$. This follows from the fact that $\exp(-\delta z^2 - iz \cdot \zeta)$ tends to $\exp(-iz \cdot \zeta)$ in $Z|_{\Omega_\varepsilon}$ as $\delta \downarrow 0$ and that the function $\exp(-\delta z^2 - iz \cdot \zeta)$ can be approximated in $Z|_{\Omega_\varepsilon}$ by functions in Z, which can be seen by Fourier transformation. Now μ is an element in the dual $Z(\mathbb{R}^n)'$ of $Z(\mathbb{R}^n)$; cf. Floret and Wloka [3, § 26, 1.2 and 1.6].

Next consider the continuous map

$$I: Z(\mathbb{R}^n) \to \prod_{k=1}^{\tilde{p}} Z(-\tilde{C}_k^*)$$

defined by restriction. A proof as in obtaining (4.8) shows that I is an injective homomorphism and that, therefore, the transposed map between the duals is surjective. Thus μ can be decomposed as

$$\mu = \sum_{k=1}^{\tilde{p}} \mu_k$$

for some $\mu_k \in Z(-\tilde{C}_k^*)'$, $k = 1, \cdots, \tilde{p}$.

For $\phi \in \mathscr{D}$ and $\eta \in \tilde{C}_k$, the limit of Riemann sums of the integral

$$\psi_\eta(z) = \int_{\mathbb{R}^n} \phi(\xi)\, e^{-iz\cdot\zeta}\, d\xi$$

converges in the space $Z(-\tilde{C}_k^*)$ and furthermore, $\psi_\eta \to \psi_0$ in $Z(-\tilde{C}_k^*)$ as $\eta \to 0$, $\eta \in \tilde{C}_k' \Subset \tilde{C}_k$. In the following we will write $\zeta_k = \xi + i\eta_k$ if $\eta_k \in \tilde{C}_k$. Now it follows that for every $\phi \in \mathscr{D}$

$$\begin{aligned}\langle g, \phi\rangle &= \sum_{k=1}^{\tilde{p}} \left\langle \mu_k, (2\pi)^{-n} \int_{\mathbb{R}^n} \phi(\xi)\, e^{-iz\cdot\xi}\, d\xi \right\rangle \\ &= \sum_{k=1}^{\tilde{p}} \lim_{\substack{\eta_k \to 0 \\ \eta_k \in \tilde{C}_k'}} \left\langle \mu_k, (2\pi)^{-n} \int_{\mathbb{R}^n} \phi(\xi)\, e^{-iz\cdot\zeta_k}\, d\xi \right\rangle \\ &= \sum_{k=1}^{\tilde{p}} \lim_{\substack{\eta_k \to 0 \\ \eta_k \in \tilde{C}_k'}} \int_{\mathbb{R}^n} (2\pi)^{-n} \langle \mu_k, e^{-iz\cdot\zeta_k}\rangle \phi(\xi)\, d\xi.\end{aligned}$$

The function

$$(4.11) \qquad h_k(\zeta)=(2\pi)^{-n}\langle \mu_k, e^{-iz\cdot\zeta}\rangle$$

is holomorphic in $\mathbb{R}^n+i\tilde{C}_k$ and for any $\varepsilon>0$ and $\tilde{C}'_k \Subset \tilde{C}_k$ h_k satisfies

$$|h_k(\zeta)| \leqq K_\varepsilon \sup_{z\in(-\tilde{C}^*_k)_\varepsilon} (1+\|x\|)^{m(\varepsilon)}|e^{-iz\cdot\zeta}|$$

$$\leqq K_\varepsilon\, e^{\varepsilon\|\xi\|} \sup_{t\geqq 0} (1+t)^{m(\varepsilon)}\, e^{-\delta(\tilde{C}'_k)\|\eta\| t}$$

$$\leqq M(\tilde{C}'_k, \varepsilon)(1+\|\eta\|^{-m(\varepsilon)})\, e^{\varepsilon\|\xi\|}, \quad \eta\in \tilde{C}'_k,$$

for some positive numbers K_ε, $m(\varepsilon)$ (depending on μ_k and ε), $\delta(\tilde{C}'_k)$ (depending on $\tilde{C}'_k$) and $M(\tilde{C}'_k, \varepsilon)$ (depending on $\tilde{C}'_k$ and ε). Hence g satisfies condition 4). □

COROLLARY 4.5. *Any analytic functional μ in Z' carried by $\mathbb{R}^n$ is the sum of boundary values in Z' of functions f_j satisfying* 3) *in Theorem* 4.4.

Remark 4.6. The abovementioned analytic representations of μ and g are in two ways not unique. Firstly, there is an arbitrariness in the number (between $n+1$ and 2^n) and the choice of the cones C_j, $j=1,\cdots,p$, or $\tilde{C}_k$, $k=1,\cdots,\tilde{p}$, as long as they satisfy (4.5) and (4.6), although (4.6) is not necessary. Secondly, if the cones $C=\bigcup_{j=1}^p C_j$ and $\tilde{C}$ are fixed, the functions f in T^C (i.e., $f=f_j$ in T^{C_j}) and h in $T^{\tilde{C}}$ are not unique. For example, instead of (4.8) g may just as well be written as

$$g_\xi=\left\{\sum_{j=1}^p g_j+\sum_{k=1}^p h^{jk}_\xi\right\}$$

where h^{jk} are arbitrary distributions with support in $C^*_j\cap C^*_k$ with the restrictions that $h^{jk}=-h^{kj}$ and that for each $\varepsilon>0$ h^{jk} can be represented as sum of weak derivatives up to order $m(\varepsilon)$ of measures $\mu^{jk}_{\alpha,\varepsilon}(\xi)$ on $C^*_j\cap C^*_k$ satisfying $\int e^{-\varepsilon\|\xi\|}|d\mu^{jk}_{\alpha,\varepsilon}(\xi)|\leqq M_\alpha(\varepsilon)$. Now $\mathcal{F}[h^{jk}]$ is the boundary value of a function holomorphic in $\mathbb{R}^n+i\,\mathrm{ch}(C_j\cup C_k)$.

Hence f' is another representation of μ if its difference with f satisfies

$$f(z)-f'(z)=\sum_{k=1}^p F_{jk}(z), \qquad y\in C_j,$$

for $j=1,\cdots,p$, where F_{jk} are arbitrary functions, holomorphic in $\mathbb{R}^n+i\mathrm{ch}(C_j\cup C_k)$ satisfying

$$|F_{jk}(z)|\leqq P(C'_j, C'_k, \varepsilon)(1+\|z\|)^{N(C'_j,C'_k,\varepsilon)}, \qquad y\in \mathrm{ch}(C'_j\cup C'_k),\quad \|y\|\geqq\varepsilon,$$

such that $F_{jk}=-F_{jk}$. Or, when $C=\bigcup\{V_{k_0}\cap\cdots\cap V_{k_{n-1}}\}$, $\{f_{k_0\cdots k_{n-1}}\}$ and $\{f'_{k_0\cdots k_{n-1}}\}$ represent the same μ if

$$f_{k_0\cdots k_{n-1}}-f'_{k_0\cdots k_{n-1}}=\sum_{j=0}^{n-1}(-1)^j F_{k_0\cdots \hat{k}_j\cdots k_{n-1}},$$

where $\hat{k}_j$ denotes that the index k_j is omitted, for arbitrary functions $F_{i_0\cdots i_{n-2}}$ holomorphic in $\mathbb{R}^n+i\{V_{i_0}\cap\cdots\cap V_{i_{n-2}}\}$ satisfying estimates as above, provided that f, f' and F are antisymmetric in their indices. Here μ is represented by the $\binom{r}{n}$ functions

$$\{f_{k_0\cdots k_{n-1}}\}_{\{k_0,\cdots,k_{n-1}\}\in\{1,\cdots,r\}}$$

if the order $k_0 \cdots k_{n-1}$ in the n-tuples $\{k_0, \cdots, k_{n-1}\}$ is such that

$$\sum_{\{k_0,\cdots,k_{n-1}\}\in\{1,\cdots,r\}} \sum_{j=0}^{n-1} (-1)^j [k_0 \cdots \hat{k}_j \cdots k_{n-1}] = 0$$

where $[i_0 \cdots i_{n-2}]$ denotes the rearrangement $i'_0 \cdots i'_{n-2}$ of $i_0 \cdots i_{n-2}$, such that $i'_0 < \cdots < i'_{n-2}$, preceded by the sign of the permutation $i_0 \cdots i_{n-2} \to i'_0 \cdots i'_{n-2}$. For example, if $r = n+1$

$$\mu = \sum_{j=1}^{n+1} (-1)^{j+1} f^*_{1\cdots\hat{j}\cdots n+1}$$

and if $r = 2^n$, $C_{\varepsilon_1\cdots\varepsilon_n} = \bigcap_{j=1}^n \{y | \varepsilon_j y_j > 0\}$ where $\varepsilon_j = \pm 1$, then

$$\mu = \sum_{\varepsilon_j \in\{-1,1\}, j=1,\cdots,n} \varepsilon_1 \cdots \varepsilon_n f^*_{\varepsilon_1\cdots\varepsilon_n};$$

cf. Martineau [5]. The representations equivalent to $h \in T^{\check{C}}$ are obtained similarly.

Remark 4.7. At a first glance Theorems 4.1, 4.2 and 4.4 2) and 3) (with (4.7) instead of (4.10)) resemble Theorems 10, 11 and 12 in § 6 of Carmichael [1]. There, too, on the one hand holomorphic functions in T^C are considered satisfying a condition similar to (4.2), namely with N constant and with P independent of r and σ in [1, formula (34)] as well as dependent on r in [1, last formula on p. 753 and formula (54), hence (49)] and on the other hand distributions U in $\mathscr{D}'$ satisfying conditions similar to ours (4.3) and (4.7), but with M_α independent of C' and r in [1, formula (48) and the last formula on p. 756] as well as dependent on C' in [1, formula (47)]. Therefore, the functions g_α in [1] are bounded or even identically zero (the function g_k in [1, last formula on p. 758]). In [1, Thms. 10, 11 and 12] finite Fourier transforms of elements $U \in \mathscr{D}'(A)$ with support in a certain, unbounded, convex set are represented as boundary values in $Z'(\overline{2\pi})$ of certain holomorphic functions. Furthermore, the distributions U are represented as weak derivatives of continuous functions on $\mathbb{R}^n$, so that also $U \in \mathscr{D}'$. However, the support of U as element of $\mathscr{D}'(A)$ is not the same as the support of U as element of $\mathscr{D}'$. According to the definition of support (see § 2) any element of $\mathscr{D}'(A)$ has a compact support. Actually, the (finite) Fourier transform of any element of $\mathscr{D}'(A)$ is the boundary value of an entire function, which result indeed is obtained in [1, Thm. 1], cf. [4, III, § 2.3]. It turns out that the proofs of Theorems 10, 11 and 12 in [1] yield a stronger result than the statements, namely they give the analytic representation in Z' of the ordinary Fourier transform of U, where U is regarded as an element of $\mathscr{D}'$. In this form [1, Thms. 10, 11 and 12] resemble our Theorems 4.1, 4.2 and 4.4 2) and 3), but although nowhere mentioned explicitly, the boundary values in [1] are always attained in S', too, and actually the theorems in [1, § 6] are particular cases of the theorems in Vladimirov [11]. Only the one dimensional "corollary" to Theorem 10 in [1, pp. 753–754] shows that boundary values in Z' are really intended. Furthermore, there is one more difference between [1, § 6] and this paper, namely, before taking the Fourier transform in [1, Thms. 11 and 12] U is reflected to $\check{U}$. This is due to the fact that the definition of Fourier transformation in [1], the one of [4], has not been "motivated by a Parseval relation." For, defining $\mathscr{F}[U]$ by requiring $\langle \mathscr{F}[U], \mathscr{F}[\phi] \rangle_{Z'} = (2\pi)^n \langle U, \phi \rangle_{\mathscr{D}'}$, $\phi \in \mathscr{D}$, one should take the complex conjugate of f in (2.1) in order to get a Parseval relation.

Example 1. Any $g \in S'$ satisfies the conditions of Theorem 4.4. Then the functions f_j in 3) of Theorem 4.4 satisfy

$$|f_j(z)| \leqq P(C')(1+\|z\|)^N (1+\|y\|^{-k}), \qquad y \in C'_j,$$

for some N and k (cf. Corollary 5.4), and a similar estimate holds for the functions h_k in 4).

Example 2. μ given by

$$\langle \mu, \psi \rangle = \oint \psi(z) \exp \frac{1}{z}\, dz$$

is an element of Z' carried by the origin. Its analytic representation is

$$\mu = \lim_{y \downarrow 0} \left\{ \exp \frac{1}{x-iy} - \exp \frac{1}{x+iy} \right\},$$

which is uniqe modulo polynomials. The inverse Fourier transform

$$\frac{1}{2\pi} \oint \exp\left(\frac{1}{z} - iz\xi\right) dz$$

is a function in $\mathbb{R}^n$, which can be continued to an entire function g with for every $\varepsilon > 0$

$$|g(\zeta)| \leqq M(\varepsilon)\, e^{\varepsilon \|\zeta\|}.$$

Here $\mathscr{F}^{-1}[\mu]$ does not belong to S'.

Example 3. μ represented as

$$\mu = \lim_{y \downarrow 0} \{ z^{i/\cos z} - \bar{z}^{i/\cos \bar{z}} \}$$

is an element of Z' carried by $(-\infty, 0] \cup \{\frac{1}{2}\pi\} \cup \{\frac{3}{2}\pi\} \cup \cdots$ and its inverse Fourier transform is a finite order distribution g in $\mathscr{D}'$, which does not belong to S'; for any $\varepsilon > 0$ g can be represented as in Theorem 4.4 2) (cf. the example at the end of § 3) or as sum of boundary values in $\mathscr{D}'$ as $\eta \to 0$, $\eta \in \tilde{C}_k' \Subset \tilde{C}_k$, of holomorphic functions h_k given by (4.11).

5. Fourier transformation as a topological isomorphism. In this section we topologize the space of distributions $g \in \mathscr{D}'$ satisfying (4.3) and the space of analytic functions f satisfying (4.2), such that the Fourier transformation $\mathscr{F}$ in Theorems 4.1 and 4.2 is a topological isomorphism. We also prove a representation theorem of such functions f and for these functions Lemma 3.3 can be improved. Most results of this section are treated earlier in de Roever [6, §§ 6 and 9], but due to the more detailed study we made in § 3, Theorem 5.5 will be proved in a shorter and less elaborate way than the corresponding theorem in [6]. For completeness, we mention two lemmas (Lemmas 5.1 and 5.2) whose proofs can be found in [6], so that here we will not give all details. We remark that the space of distributions g given here has a more simple form than in [6]. In de Roever [7], [8] the theorems of this section are used to derive the Fourier transformation between functions f of exponential type in both $\|x\|$ and $\|y\|$, holomorphic in tubular radial domains, and certain spaces of analytic functionals with unbounded, convex, carrier. In general, these functions do not have distributional boundary values on the distinguished boundary, but as a particular case the analytic representation is obtained of distributions in $\mathscr{D}'$ being the Fourier transform of elements in Z' carried by certain, unbounded, convex sets in $\mathbb{C}^n$. In this form the theorems of [7, Thm. 6.1] and [8, III] are opposite to the theorems of this section.

First we remark that the correspondence between the exponential type $b\|y\|$ of f in Theorem 4.1 and the set (4.1) $\{\xi | -y \cdot \xi \leqq b, y \in \operatorname{pr} C\}$ can be generalized to exponential types varying with the direction of y and arbitrary convex sets. Let b be a convex function of $y \in C$ homogeneous of degree one, where C is an open convex cone in $\mathbb{R}^n$. This means that $b(y)$ is determined by its value on pr C:

$$b(y) = \|y\| b\left(\frac{y}{\|y\|}\right).$$

The convex open cone C and the convex homogeneous function b on C determine a closed convex set $U = U(b, C)$ in $\mathbb{R}^n$ by

$$U(b, C) \stackrel{\text{def}}{=} \{\xi | -y \cdot \xi \leqq b(y), y \in C\}. \tag{5.1}$$

If b can be continuously continued to pr $\bar{C}$, then C and $\bar{C}$ determine the same convex set $U(b, C)$.

Conversely, each closed convex set U in $\mathbb{R}^n$ determines an open (possibly in some linear subspace of $\mathbb{R}^n$), convex cone C in $\mathbb{R}^n$ and a convex, homogeneous function b on C by: let for $y \in \mathbb{R}^n$ and for some real number α $H(y, \alpha)$ be the affine half space in $\mathbb{R}^n$

$$H(y, \alpha) = \{\xi | -y \cdot \xi \leqq \alpha\};$$

then C is the interior (possibly in some linear subspace of $\mathbb{R}^n$) of the set of all $y \in \mathbb{R}^n$ such that $U \subset H(y, \alpha)$ for a real number α depending on y and

$$b(y) = \sup_{\xi \in U} -y \cdot \xi. \tag{5.2}$$

C is open in $\mathbb{R}^n$ (hence, C is not contained in a proper linear subspace) if and only if U does not contain a straight line. Note that $b(y)$ might not be positive for all $y \in$ pr C, in which case U determined by (5.1) does not contain the origin.

Secondly we discuss representations of distributions g with support in a closed set U as sum of weak derivatives of measures. For arbitrary sets U such a representation is not always possible, because U has to satisfy certain properties, see Whitney [13] and more generally Schwartz [9] or Vladimirov [11], [12]. Here we only need that it is sufficient if U is the closure of a convex, open set. It is shown in Whitney [14] that a C^∞-function ϕ on the closed convex set U (see Whitney [13]), whose derivatives are uniformly continuous and bounded on U, can be extended to a C^∞-function on an ε-neighborhood of U, which is bounded there. Hence ϕ can be extended to a C^∞-function $\tilde{\phi}$ on $\mathbb{R}^n$ which, together with its derivatives, is bounded. Moreover, it follows from the construction of $\tilde{\phi}$ in [14] that, if $D^\alpha \phi(\xi) \to 0$ as $\xi \to \infty$ in U (hence then $D^\alpha \phi$ is uniformly continuous in U), this also holds for $\tilde{\phi}$ as $\xi \to \infty$ in $\mathbb{R}^n$.

Let $\alpha(m)$ and $\beta(m)$ be two nondecreasing sequences, where for every m at least one of the inequalities $\alpha(m+1) > \alpha(m)$ or $\beta(m+1) > \beta(m)$ holds and let $M_m(\xi)$ be a positive C^∞-function on $\mathbb{R}^n$ which outside the unit ball equals

$$(1 + \|\xi\|)^{\alpha(m)} e^{\beta(m)\|\xi\|}.$$

For any closed set V the norms

$$\sup_{\substack{\xi \in V \\ |\alpha| \leqq m}} |D^\alpha M_m(\xi) \phi(\xi)| \quad \text{and} \quad \sup_{\substack{\xi \in V \\ |\alpha| \leqq m}} M_m(\xi) |D^\alpha \phi(\xi)|$$

are equivalent, and also it is equivalent to assert

$$\lim_{\substack{\xi \to \infty \\ \xi \in V}} D^\alpha M_m(\xi) \phi(\xi) = 0, \qquad |\alpha| \leqq m,$$

or

$$\lim_{\substack{\xi\to\infty\\ \xi\in V}} M_m(\xi)D^\alpha\phi(\xi)=0, \qquad |\alpha|\leqq m.$$

Let us denote by $W^m_{\infty,0}(M_m;\,V)$ the Banach space of C^m-functions ϕ on the closure V of an open, convex set (in the sense of Whitney [13]) with the norm

$$\|\phi\|_m=\sup_{\substack{\xi\in V\\ |\alpha|\leqq m}} M_m(\xi)|D^\alpha\phi(\xi)|$$

and with

$$\lim_{\substack{\xi\to\infty\\ \xi\in V}} M_m(\xi)D^\alpha\phi(\xi)=0, \qquad |\alpha|\leqq m,$$

(cf. Wloka [15]) and by $W(V)$ the Fréchet space

$$W(V)=\operatorname{proj}\lim_{m\to\infty} W^m_{\infty,0}(M_m;\,V).$$

Then the restriction map I from $W(\mathbb{R}^n)$ into $W(U)$ is surjective; cf. Vladimirov [12] in case $\beta(m)\equiv 0$. According to Treves [10, Thm. 37.2] the transposed map I' from the dual $W(U)'$ of $W(U)$ into the dual $W(\mathbb{R}^n)'$ of $W(\mathbb{R}^n)$ is injective and has weakly closed range. Therefore, $W(U)'$ can be identified (by means of I') with the subspace W'_U of $W(\mathbb{R}^n)'$ consisting of the elements with support in U. Indeed, W'_U, by the definition of support (see § 2) vanishing on the space of all $\phi\in W(\mathbb{R}^n)$ with support in U^c, also vanishes on the closure of this space, which is just Ker I. Then according to Treves [10, Prop. 35.4] W'_U is the weak closure of Im I', and since this is already weakly closed, $I'(W(U)')=W'_U$.

Furthermore, we may conclude that $W(U)'$ is a closed linear subspace of $W(\mathbb{R}^n)'$. For it follows from Wloka [15] that the identity map from $W^{m+1}_{\infty,0}(M_{m+1};\mathbb{R}^n)$ into $W^m_{\infty,0}(M_m;\mathbb{R}^n)$ is compact, hence that $W(\mathbb{R}^n)$ is an $F\bar{S}$-space; see Floret and Wloka [3] ($W(\mathbb{R}^n)$ is even a nuclear $F\bar{S}$-space). Hence $W(\mathbb{R}^n)'$ can be written as inductive limit (LS-space) and it is reflexive. Therefore, Im I' is even strongly closed. Now the following natural embedding maps are bijective and continuous:

$$\operatorname{ind}\lim_{m\to\infty} W^m_{\infty,0}(M_m;\,U)'\to W(U)'\xrightarrow{I'} W'_U\subset W(\mathbb{R}^n)'.$$

By a property of LS-spaces (Floret and Wloka [3, 25.1]) the closed sets of the first space are closed in W'_U, where $W'_U=\operatorname{Im} I'$ is regarded as a closed subspace of $W(\mathbb{R}^n)'$. Therefore, the three spaces are equal also as topological spaces. Thus $W(U)'$ is an LS-space, hence $W(U)$ is an $F\bar{S}$-space, and $W(U)'$ is a closed linear subspace of $W(\mathbb{R}^n)'$.

Finally, by Riesz' representation theorem distributions $g\in W(U)'$ can be represented as sum of weak derivatives of measures μ_α on U such that

$$g=\sum_{|\alpha|\leqq m} D^\alpha\mu_\alpha, \qquad \int_U \frac{|d\mu_\alpha(\xi)|}{M_k(\xi)}<\infty, \qquad |\alpha|\leqq m,$$

where m and k depend on g. For this reason we required that $M_m(\xi)D^\alpha\phi(\xi)\to 0$ as $\xi\to\infty$ in U, for, in that case the measures μ_α, according to Riesz' theorem being defined on a compactification of U, are still concentrated on U.

We now describe the space of the distributions g of Theorem 4.2. In the remaining C will be an open, convex, cone in $\mathbb{R}^n$. Theorem 4.2 holds for convex homogeneous functions $b(y)$ instead of $b\|y\|$ as well. Let

$$S^k(b, C) \overset{\text{def}}{=} \operatorname*{proj\,lim}_{m\to\infty} W^m_{\infty,0}((1+\|\xi\|)^m \exp((1/k)\|\xi\|); U(b, C)), \tag{5.3}$$

where $U(b, C)$ is given by (5.1). For $p > k$ the identity map maps $S^k(b, C)$ continuously into $S^p(b, C)$, hence the strong dual $S^p(b, C)'$ of $S^p(b, C)$ into $S^k(b, C)'$. Now define the space

$$S^*(b, C)' \overset{\text{def}}{=} \operatorname*{proj\,lim}_{k\to\infty} S^k(b, C)'.$$

Deleting $(1+\|\xi\|)^m$ from the weight functions in (5.3) would yield the same space $S^*(b, C)'$, but $S^k(b, C)$ in the form (5.3) is an $F\bar{S}$-space. We choose an increasing sequence $\{C_k\}_{k=1}^\infty$ of convex subcones of C exhausting C, such that, when $\delta_k > 0$ is a number with for $y \in C_k$ and $\xi \in C^*_{k+1}$

$$y \cdot \xi \geqq \delta_k \|y\| \, \|\xi\| \qquad \text{(cf. (4.9))}, \tag{5.4}$$

then $\{1/k\delta_k\}_{k=1}^\infty$ is a decreasing sequence of positive numbers. In view of the fact that for any k the set $(C^*_{k+1})^c \cap U(b, C)$ is compact (see the proof of Lemma 5.1) the distributions g in $S^*(b, C)'$ are just the $g \in \mathcal{D}'$ of Theorem 4.2.

LEMMA 5.1. *For all k there is a $p > k$ such that for any $z = x + iy$ with $y \in C_k$ and $\|y\| > 1/k$*

$$e^{iz\cdot\xi} \in S^p(b, C)_\xi.$$

Proof. See de Roever [6, Lemma 9.1]. In [6, Lemma 6.3] it is shown that $\xi \in (C^*_{k+1})^c \cap U(b, C)$ satisfies $\|\xi\| \leqq d_k$ for some positive number d_k depending on k.

Now let $p > k/\delta_k$; then using (5.4) and (5.2) we find for $y \in C_k$, $\|y\| > 1/k$ and every m

$$\begin{aligned}
&\sup_{\substack{\xi\in U(b,C)\\ |\alpha|\leqq m}} (1+\|\xi\|)^m \exp\left(\frac{1}{p}\|\xi\|\right) |D^\alpha e^{iz\cdot\xi}| \\
&\leqq \sup_{\substack{\xi\in C^*_{k+1}\\ |\alpha|\leqq m}} (1+\|\xi\|)^m |z^\alpha| \exp\left(\frac{1}{p} - \delta_k \frac{1}{k}\right)\|\xi\| \\
&\qquad + \sup_{|\alpha|\leqq m} (1+d_k)^m |z^\alpha| \exp\left(\frac{1}{p} d_k + b(y)\right) \\
&\leqq M_k (1+\|z\|)^m e^{b(y)} \sup_{t\geqq 0} (1+t)^m \exp -\left(\frac{1}{k}\delta_k - \frac{1}{p}\right) t.
\end{aligned}$$

The lemma follows from the fact that for all $t \geqq 0$

$$(1+t)^m \exp -\delta t \leqq K_m 1/\delta^m$$

for some constant K_m depending on m. □

As an element of S' the Fourier transform of $e^{-y\cdot\xi} g_\xi$ with $g \in S^*(b, C)'$ is known. We can now formulate a simple representation of this Fourier transform.

LEMMA 5.2. *For any $y \in C$ and $g \in S^*(b, C)'$*

$$\mathcal{F}[e^{-y\cdot\xi} g_\xi](x) = \langle g, e^{iz\cdot\xi} \rangle. \tag{5.5}$$

Proof. See de Roever [6, Lemma 9.2]. Let us first assume that moreover $g \in S'$. Let α be a C^∞-function with support in $U(b+1, C)$, equal to 1 in $U(b, C)$, such that α is a multiplier in S' (here $(b+1)(y) = b(y) + \|y\|$). Then $\alpha(\xi)\, e^{-y\cdot\xi} \in S_\xi$ for every $y \in C$. Essentially by a change of order of integration it is shown in [6, Lemma 6.4] that

$$\mathscr{F}[e^{-y\cdot\xi} g_\xi](x) = \langle g_\xi, \alpha(\xi)\, e^{iz\cdot\xi}\rangle_{S'}.$$

Now we take $g \in S^*(b, C)'$ and $y \in C$. Let $y_0 \in C$ be such that $y - y_0 \in C$ and let p be such that for ξ in $U(b, C)$ outside a compact set $-y_0 \cdot \xi \leqq (-1/p)\|\xi\|$. Then multiplication by $\exp -y_0 \cdot \xi$ and restriction to $U(b, C)$ is a continuous map from S into $S^p(b, C)$, so its transpose is continuous from $S^p(b, C)'$ into S'. Hence $e^{-y_0\cdot\xi} g_\xi \in S'$ and according to the above

$$\begin{aligned}\mathscr{F}[e^{-(y-y_0)\cdot\xi}\, e^{-y_0\cdot\xi} g_\xi](x) &= \langle e^{-y_0\cdot\xi} g_\xi, \alpha(\xi)\, e^{i(z-iy_0)\cdot\xi}\rangle_{S'}\\ &= \langle g_\xi, e^{-y_0\cdot\xi}\alpha(\xi)\, e^{i(z-iy_0)\cdot\xi}|_{U(b,C)}\rangle_{S^p(b,C)'} = \langle g_\xi, e^{iz\cdot\xi}\rangle.\end{aligned}$$

Theorem 4.1 also holds for convex homogeneous functions $b(y)$ instead of $b\|y\|$, if C is convex. Therefore, we get a representation of the function f of Theorem 4.1 (cf. Carmichael [1, Thm. 13]). Let f be holomorphic in T^C and satisfy

$$(5.6)\qquad |f(z)| \leqq P(C', r, \sigma)(1+\|z\|)^{N(C',r)} \exp\{b(y) + \sigma\|y\|\}, \qquad z \in T^{C'}, \quad \|y\| \geqq r,$$

for all $C' \Subset C$, $r > 0$ and $\sigma > 0$.

COROLLARY 5.3. *For any function f that satisfies* (5.6) *there is a distribution $g \in S^*(b, C)'$ such that*

$$(5.7)\qquad f(z) = \langle g_\xi, e^{iz\cdot\xi}\rangle.$$

With this representation Lemma 3.3 can be improved so that $m(C')$ in (3.8) no longer depends on C'.

COROLLARY 5.4. *Let the boundary value f^* in Z' of a function f satisfying* (5.6) *belong to S'. Then f attains this boundary value already in S' as $y \to 0$, $y \in C'$ and f satisfies the stronger condition*

$$(5.8)\qquad |f(z)| \leqq P(C')(1+\|z\|)^m (1+\|y\|^{-k}) \exp b(y), \qquad y \in C',$$

for every $C' \Subset C$ and some m and k.

Proof. In view of (5.7) it is sufficient to represent $g = \mathscr{F}^{-1}[f^*]$ as sum of weak derivatives of measures in $U(b, C)$ and to estimate

$$\sup_{\substack{\xi \in U(b,C)\\ |\alpha| \leqq m}} (1+\|\xi\|)^k |D^\alpha e^{iz\cdot\xi}|$$

as in the proof of Lemma 5.1. □

We now define a topology on the space $H^*(b, C)$ of functions f satisfying (5.6) by

$$(5.9)\qquad H^*(b, C) \overset{\text{def}}{=} \operatorname*{proj\,lim}_{k\to\infty}\, \operatorname*{ind\,lim}_{m\to\infty} A_\infty\Big(\frac{\exp - b(y)}{(1+\|z\|)^m}; \mathbb{R}^n + iC(k)\Big),$$

where $A_\infty(M(z); \Omega)$ denotes the Banach space of holomorphic functions in Ω with finite sup norm $\sup_{z\in\Omega} M(z)|f(z)|$ and where $C(k) = C_k \cap \{y \mid \|y\| > 1/k\}$. Here the continuous maps from

$$H^p_m \overset{\text{def}}{=} A_\infty\Big(\frac{\exp - b(y)}{(1+\|z\|)^m}; \mathbb{R}^n + iC(p)\Big)$$

into H_l^k, $l \geqq m$, $p \geqq k$, are the natural injections. By changing this representation of the space $H^*(b, C)$ somewhat, one can see that $H^*(b, C)$, just as $S^*(b, C)'$, is the projective limit of nuclear LS-spaces: let for each k $\{C_{k+1/m}\}_{m=1}^{\infty}$ be a decreasing sequence of convex, relatively compact subcones of C_{k+1} with intersection $\overline{C_k}\backslash\{0\}$ and let $C(k+1/m) = C_{k+1/m} \cap \{y \mid \|y\| > k+1/m\}$. Then also

$$H^*(b, C) = \operatorname{proj\,lim}_{k\to\infty} \operatorname{ind\,lim}_{m\to\infty} H_m^{p+1/m}$$

and from the compact embedding theorems between A-spaces in Wloka [15, Thm. 2, § 4.2, where the condition $d(S_n, CS_{n+1}) > 0$ may be replaced by $\overline{S_n} \subset G_1$] follows the abovementioned property.

The following theorem gives the Fourier transformation in Theorems 4.1 and 4.2 as an isomorphism (cf. de Roever [6, Thm. 9.1]).

THEOREM 5.5. *The Fourier transformation* $\mathcal{F}: S^*(b, C)' \to H^*(b, C)$ *given by* $\mathcal{F}(g)(z) = \langle g_\xi, e^{iz\cdot\xi}\rangle$ *for* $g \in S^*(b, C)'$ *is a topological isomorphism.*

Proof. In the proof of Lemma 5.2 it is shown that for $g \in S^*(b, C)'$ and $y_0 \in C$, $e^{-y_0\cdot\xi} g_\xi$ indeed belongs to S', so that the Fourier transformation (5.5) is 1–1. Hence $\mathcal{F}$ is an injective map from $S^*(b, C)'$ into $H^*(b, C)$ according to Theorem 4.2. Theorem 4.1 says that this map is moreover surjective. In order to prove the continuity of $\mathcal{F}$ it is sufficient to show that for each k there is a p such that $\mathcal{F}$ is a bounded map from $S^p(b, C)'$ into

$$H^k(b, C) \overset{\text{def}}{=} \operatorname{ind\,lim}_{m\to\infty} H_m^k,$$

because as an LS-space $S^p(b, C)'$ is bornologic; see Floret and Wloka [3]. So, let B be a bounded set in $S^p(b, C)'$ where p is still to be chosen. This means that for some m B is bounded in the strong dual of $W_{\infty,0}^m(M_m^p; U(b, C))$, where $M_m^p(\xi) = (1+\|\xi\|)^m \exp((1/p)\|\xi\|)$. Hence there is a K such that for all $\phi \in W_{\infty,0}^m(M_m^p; U(b, C))$ and $g \in B$

$$|\langle g, \phi\rangle| \leqq K \sup_{\substack{\xi\in U(b,C)\\ |\alpha|\leqq m}} M_m^p(\xi)|D^\alpha\phi(\xi)|.$$

Now we choose $p > k$ as in Lemma 5.1 and replacing $\phi(\xi)$ by $\exp(iz\cdot\xi)$ in the above estimate yields for the images $f = \mathcal{F}(g)$

$$|f(z)| \leqq KM_{k,m}(1+\|z\|)^m e^{b(y)}, \qquad y \in C_k, \quad \|y\| > 1/k.$$

Hence $\mathcal{F}(B)$ is bounded in H_m^k, thus bounded in $H^k(b, C)$.

Next we prove the continuity of $\mathcal{F}^{-1}$. Again it would be sufficient to show that for each k there is a p such that $\mathcal{F}^{-1}$ is a bounded map from the LS-space $\operatorname{ind}_{m\to\infty}\lim H_m^{p+1/m}$ into $S^k(b, C)'$. So, let us start with a bounded set in $H_m^{p+1/m}$ for some m. This set is certainly bounded in H_m^p. However, the elements of a bounded set A in H_m^p are holomorphic in $\mathbb{R}^n + iC(p)$, so that we cannot expect that $\mathcal{F}^{-1}(A) \subset S^k(b, C)'$.

Let $p > k$ and let $y_0 \in C_p$ be such that

$$\frac{1}{p} \leqq \|y_0\| \leqq \frac{1}{k}.$$

Then $f(z+iy_0)$ is holomorphic in $\mathbb{R}^n+iC_p$ if $f\in A$, and it satisfies there

$$|f(z+iy_0)|\leqq M(1+\|z\|)^m \exp b(y+y_0)\leqq M'(1+\|z\|)^m \exp b(y),$$

because $b(y)$ is homogeneous and convex: $\frac{1}{2}b(y+y_0)=b(\frac{1}{2}y+\frac{1}{2}y_0)\leqq\frac{1}{2}b(y)+\frac{1}{2}b(y_0)$. Hence the set $B'=\{g'|g'=\mathscr{F}^{-1}[f(x+iy_0)], f\in A\}$ is a bounded set in S' and every $g'\in B'$ has its support in the set $U(b, C_p)$. Since $y_0\cdot\xi\leqq(1/k)\|\xi\|$, multiplication by $\exp(y_0\cdot\xi)$ maps B' into a bounded set B in $S^k(b, C_p)'$. According to Corollary 3.2 and (5.5), for $f\in A$ and for all $y\in\{y|y+y_0\in C(p)\}$ with $\exp[i(z+iy_0)\cdot\xi]\in S^k(b, C_p)$, we have

$$f(z+iy_0)=\mathscr{F}[e^{-y\cdot\xi}g'_\xi]=\mathscr{F}[e^{-(y+y_0)\cdot\xi}e^{y_0\cdot\xi}g'_\xi]=\langle g_\xi, e^{i(z+iy_0)\cdot\xi}\rangle, \tag{5.10}$$

for some $g\in B$. As in (3.4) g is independent of y_0, so that $\mathscr{F}^{-1}(f)=g$. Hence $\mathscr{F}^{-1}(A)$ is bounded in $S^k(b, C_p)'$. If f also belongs to H_l^p for a larger p, $l\geqq m$, then still we would have found the same g. Therefore, $\mathscr{F}^{-1}$ is a continuous map from $H^*(b, C)$ into $S^k(b, C_p)'$ for any p with the same image in every space $S^k(b, C_p)'$, $p=1, 2, \cdots$. Thus $\mathscr{F}^{-1}$ is continuous from $H^*(b, C)$ into $\text{proj}_{p\to\infty}\lim S^k(b, C_p)$, which equals $S^k(b, C)'$ because $S^k(b, C_{p+1})'$ is a closed linear subspace of $S^k(b, C_p)'$. Hence $\mathscr{F}^{-1}$ is continuous from $H^*(b, C)$ into $S^k(b, C)'$ for all k, and since in (5.10) g is also independent of k, it follows that $\mathscr{F}^{-1}$ is a continuous map from $H^*(b, C)$ into $S^*(b, C)'$. □

Similarly, when the boundary values of the functions f exist in S', we can topologize the spaces of these functions and of their inverse Fourier transforms, so that the Fourier transformation is a topological isomorphism. For that purpose, let

$$S(b, C) \overset{\text{def}}{=} \operatorname*{proj\,lim}_{m\to\infty} W^m_{\infty,0}((1+\|\xi\|)^m; U(b, C))$$

and let

$$H(b, C) \overset{\text{def}}{=} \operatorname*{ind\,lim}_{m\to\infty} \ \operatorname*{proj\,lim}_{k\to\infty} A_\infty\Big(\frac{\exp[-b(y)]}{(1+\|z\|)^m(1+\|y\|^{-m})}; \mathbb{R}^n+iC_k\Big).$$

$S(b, C)$ is an $F\bar{S}$-space and as a consequence of the following theorem $H(b, C)$ is an LS-space. Also here for any $z\in T^C$ $\exp(iz\cdot\xi)\in S(b, C)_\xi$ and Lemma 5.2 holds for $g\in S(b, C)'$. Similarly to Theorem 5.5 with the aid of Vladimirov [11, Thm. 26.3 and 26.4, Thm. 2] one can prove the following theorem, which gives the Fourier transformation of [11, 26.4, Thm. 2] as a topological isomorphism (cf. de Roever [6, Thm. 6.1]).

THEOREM 5.6. *The Fourier transformation $\mathscr{F}$: $S(b, C)'\to H(b, C)$ given by $\mathscr{F}(g)(z)=\langle g_\xi, e^{iz\cdot\xi}\rangle$ for $g\in S(b, C)'$ is a topological isomorphism.*

Remark 5.7. At the beginning of this section it is shown that $S(b, C)'$ is a closed linear subspace of S'. However, $S^*(b, C)'$ as a subset of $\mathscr{D}'$ carries a finer topology than the one induced by $\mathscr{D}'$. For, let $n=1$, $C=\{y|y>0\}$ and $b=0$, then the function $\theta(\xi)e^\xi$ belongs to $\mathscr{D}'$, but not to $S^*(b, C)'$, because $\theta(\xi)e^{(1-y)\xi}\notin S'$ if $0<y<1$ (here $\theta(\xi)=1$ if $\xi>0$ and $\theta(\xi)=0$ if $\xi<0$). Furthermore $\theta(\xi)\sum_{k=0}^N(1/k!)\xi^k$ belongs to $S^*(b, C)'$ for every N and the limit for $N\to\infty$ converges in $\mathscr{D}'$ to $\theta(\xi)e^\xi$, hence it does not converge in $S^*(b, C)'$, which as projective limit of complete spaces is itself complete.

We end this paper with a last striking property of functions in $H^*(b, c)$ or $H(b, C)$, when the function b is moreover uniformly continuous in C. Let $w\in\text{pr}\,C$; then $w\in\text{pr}\,C'$ for some $C'\Subset C$. Since for $C'\Subset C''\Subset C$ $U(b, C)\cap C''_*$ is bounded and since for $\xi \in C''^*$

$$\delta\|\xi\| \leqq w \cdot \xi \leqq \|\xi\|$$

for some $\delta > 0$, in (5.3) the weight functions $\exp((1/k)\|\xi\|)$ in the definition of $S^*(b, C)'$ may be replaced by $\exp((1/k)w \cdot \xi)$. Then a function $f \in H^*(b, C)$ satisfies for each $\varepsilon > 0$ and $z \in \mathbb{R}^n + i\{\varepsilon w + C\}$

$$\begin{aligned} |f(z)| &\leqq |\langle g_\xi, e^{iz\cdot\xi}\rangle| \\ &\leqq K''_\varepsilon(1+\|z\|)^{m(\varepsilon)} \sup_{\xi \in U(b,C)} (1+\|\xi\|)^{m(\varepsilon)} e^{(1/2)\varepsilon w\cdot\xi - y\cdot\xi} \\ &\leqq K'_\varepsilon(1+\|z\|)^{m(\varepsilon)} \sup_{\xi \in U(b,C)} \exp(\varepsilon w \cdot \xi - y \cdot \xi) \leqq K'_\varepsilon(1+\|z\|)^{m(\varepsilon)} \exp b(y - \varepsilon w) \\ &\leqq K_\varepsilon(1+\|z\|)^{m(\varepsilon)} e^{b(y)}, \end{aligned}$$

because b is uniformly continuous in C. Since this property for functions $f \in H^*(b, C)$ is not true for general convex homogeneous functions b, we see that it was right to divide C into $\bigcup_{k=1}^{\infty} C(k)$ instead of $\bigcup_{\varepsilon>0} \{\varepsilon w + C\}$ in the definition (5.9) of $H^*(b, C)$. A similar property holds for functions f in $H(b, C)$.

For example, we may take b constant on pr C, i.e., $b(y) = b\|y\|$ where now b is a number. Then we have obtained the following corollary.

COROLLARY 5.8. *If a holomorphic function f in T^C, C open and convex, satisfies* (4.2), *then it also satisfies, for every $\varepsilon > 0$,*

$$|f(z)| \leqq P(\varepsilon)(1+\|z\|)^{N(\varepsilon)} \exp(b\|y\|), \qquad y \in \varepsilon w + C,$$

for some fixed $w \in$ pr C. If f satisfies (5.8), *where $b(y) = b\|y\|$ for some number b, then f also satisfies for some $m' \geqq m$ and every $\varepsilon > 0$*

$$|f(z)| \leqq P(\varepsilon)(1+\|z\|)^{m'} \exp(b\|y\|), \qquad y \in \varepsilon w + C.$$

REFERENCES

[1] R. D. CARMICHAEL, *Analytic representation of the distributional finite Fourier transform*, this Journal, 5 (1974), pp. 737–761.

[2] L. EHRENPREIS, *Fourier Analysis in Several Complex Variables*, John Wiley, New York, 1970.

[3] K. FLORET AND J. WLOKA, *Einführung in die Theorie der lokalkonvexen Räume*, Lecture Notes in Mathematics, no. 56, Springer-Verlag, Berlin, 1968.

[4] I. M. GEL'FAND AND G. E. SHILOV, *Generalized Functions*, vol. 2, Academic Press, New York, 1968.

[5] A. MARTINEAU, *Distributions et valeurs au bord des fonctions holomorphes*, Theory of distributions, Proc. of the Internat. Summer Inst. Lisbon, 1964, Inst. Gulbenkian Ciência, Lisboa, 1964, pp. 193–326.

[6] J. W. DE ROEVER, *Fourier transforms of holomorphic functions and application to Newton interpolation series, I*, Afdeling Toegepaste Wiskunde, Rep. TW 142, Math. Centrum, Amsterdam, 1974.

[7] ———, *Fourier transforms of holomorphic functions and application to Newton interpolation series, II*, Afdeling Toegepaste Wiskunde, Rep. TW 148, Math. Centrum, Amsterdam, 1975.

[8] ———, *Complex Fourier transformation and analytic functionals with unbounded carriers*, MC Tracts 89, Math. Centrum, Amsterdam, 1977.

[9] L. SCHWARTZ, *Théorie des Distributions*, Hermann, Paris, 1966.

[10] F. TREVES, *Topological Vector Spaces, Distributions and Kernels*, Pure and Applied Mathematics Series, no. 25, Academic Press, N.Y., 1967.

[11] V. S. VLADIMIROV, *Methods of the Theory of Functions of Many Complex Variables*, MIT Press, Cambridge, MA, 1966.

[12] ———, *Functions which are holomorphic in tubular cones*, Izv. Akad. Nauk. SSSR Ser. Mat., 27 (1963), pp. 75–100.

[13] H. WHITNEY, *Functions differentiable on the boundary of regions*, Ann. of Math., 35 (1934), pp. 482–485.

[14] ———, *Analytic extensions of differentiable functions defined in closed sets*, Trans. Amer. Math. Soc., 36 (1934), pp. 63–89.

[15] J. WLOKA, *Grundräume und verallgemeinerte Funktionen*, Lecture Notes in Mathematics, no. 82, Springer-Verlag, Berlin, 1969.

SIAM. J. MATH. ANAL.
Vol. 9, No. 6, December 1978
© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0003 $01.00/0

ON THE SOLUTIONS OF A CLASS OF NONLINEAR STURM–LIOUVILLE PROBLEMS*

P. DE MOTTONI† AND A. TESEI†

Abstract. A class of Sturm–Liouville problems with monomial nonlinearities is studied in a constructive way by explicit integration. Results are obtained concerning the existence and uniqueness of the solutions with a prescribed number of nodes, the location of the nodes, the behavior of the maximum norm of the solutions as functions of a distinguished parameter (bifurcation parameter).

1. Introduction. We want to study the nonlinear Sturm–Liouville problem

$$u''+\lambda u-u^k=0 \quad \text{in } (0, T) \qquad (1)$$
$$u(0)=u(T)=0$$

where $T\in\mathbb{R}^+$, $k\in\mathbb{N}$, $k\geqq 2$; namely we are interested in finding the values $\lambda\in\mathbb{R}$ such that the problem (1) admits a real nontrivial solution[1] u, and, in particular, the values giving rise to a solution with m intermediate zeros (nodes). We shall be concerned as well with a finer analysis of the location in $(0, T)$ of the nodes, and with the relationship between the parameter λ (the "eigenvalue") and the (supremum) norm of the corresponding solutions.

Corresponding results will also be given for

$$u''+\lambda u+u^k=0 \quad \text{in} \quad (0, T) \qquad (1')$$
$$u(0)=u(T)=0$$

whose investigation can be reduced to that of (1) by simple arguments.

As it is well known, nonlinear problems of Sturm–Liouville type have been often considered in the framework of bifurcation theory as an application of general theorems [1], [2]. In the present paper we stipulate to follow a constructive approach, which only involves elementary techniques. Although this procedure cannot be used for a wide class of problems,[2] performing explicit calculations in a nontrivial instance not only gives a concrete illustration of the general results, but also shows very neatly the structure of the bifurcated solutions, and provides a deeper insight into their properties.

The paper is organized as follows: After a summary of the results (§ 2), we shall perform a phase plane analysis of (1), by which a first characterization of the appearence of solutions will be obtained (§ 3). The proofs are contained in § 4.

2. Summary of the results. Let us denote by μ_m the eigenvalues of the linear problem associated with (1), namely $\mu_m=((m+1)\cdot\pi/T)^2$ $(m\geqq 0)$. In the sequel we shall prove the following results:

THEOREM 1. *When k is odd, a unique couple $(u, -u)$ of solutions of* (1) *with m nodes exists if $\lambda>\mu_m$ $(m\geqq 0)$. If $\lambda\leqq\mu_m$, no solution with m nodes exists.*

THEOREM 2. *When k is even, the following situation prevails*:

(i) *A unique positive solution exists if and only if $\lambda<\mu_0$; a unique negative solution exists if and only if $\lambda>\mu_0$;*

* Received by the editors February 25, 1976, and in revised form March 8, 1977.
† Istituto per le Applicazioni del Calcolo (IAC), "Mauro Picone", Rome, Italy.
[1] Notice that any weak solution of (1) is a classical one: hence we may confine ourselves to consider smooth solutions.
[2] See, however, reference [3].

(ii) *A unique couple of solutions* $(u'(0) \gtrless 0)$ *with* $2m+1$ *nodes exist if* $\lambda > \mu_{2m+1}(m \geqq 0)$. *If* $\lambda < \mu_{2m+1}$, *no solution with* $2m+1$ *nodes exists.*

(iii) *A unique solution with* $2m$ *nodes such that* $u'(0) > 0$ *exists if and only if* $\lambda > \mu_{2m}$ $(m \geqq 1)$;

(iv) *As for the solutions with* $2m$ *nodes and* $u'(0) < 0$, $(m \geqq 1)$, *there exists a decreasing positive sequence* $\{\tau_m\}$ *such that: no solution of this kind exists for* $\lambda < \mu_{2m} - \tau_m$; *one solution exists for* $\lambda = \mu_{2m} - \tau_m$ *and for* $\lambda > \mu_{2m}$; *two solutions exist for* $\lambda \in (\mu_{2m} - \tau_m, \mu_{2m})$.

Additional results, concerning the behavior of the supremum norm of the solutions as a function of λ are summarized in the following propositions:

PROPOSITION 1. *The supremum norm of any solution of* (1) *is for large* $|\lambda|$ *an increasing function of* $|\lambda|$, *which behaves as* $|\lambda|^{1/(k-1)}$. *The above monotonicity property holds for any* λ *in the case k odd and, in the case k even, for the solutions of items* (i), (ii), (iii) *of Theorem* 2.

PROPOSITION 2. *For k odd, the supremum norm of any solution of* (1), *viewed as a function of* λ, *has an infinite right derivative at* $\lambda = \mu_m$ $(m \geqq 0)$. *For* $k = 2$, *the supremum norm of the positive solution has a finite right derivative at* $\lambda = \mu_0$, *which coincides with the left derivative of the norm of the negative solution. For k even and different from* 2, *such derivatives are infinite.*

Due to the above results, we can draw the bifurcation diagrams for our problem, which are depicted in Figs. 1, 2.

As for the problem (1′), in the case k even, it is reduced to (1) via the transformation $u \to -u$: thus the bifurcation diagram is obtained from Fig. 2 by reflection with respect to the λ-axis. In the case k odd, we shall prove that the bifurcation diagram is as in Fig. 3.

It will become transparent in the following that, in the case k odd, the nodes of any solution of (1) are equally spaced in $[0, T]$. The same is no longer true when k is even. A precise result is expressed by the following proposition:

PROPOSITION 3. *Consider, for k even, a solution u of* (1) *with* $2m+1$ *nodes* $(m \geqq 0)$. *Then the length of the subintervals of* $[0, T]$ *where u is positive is strictly larger*

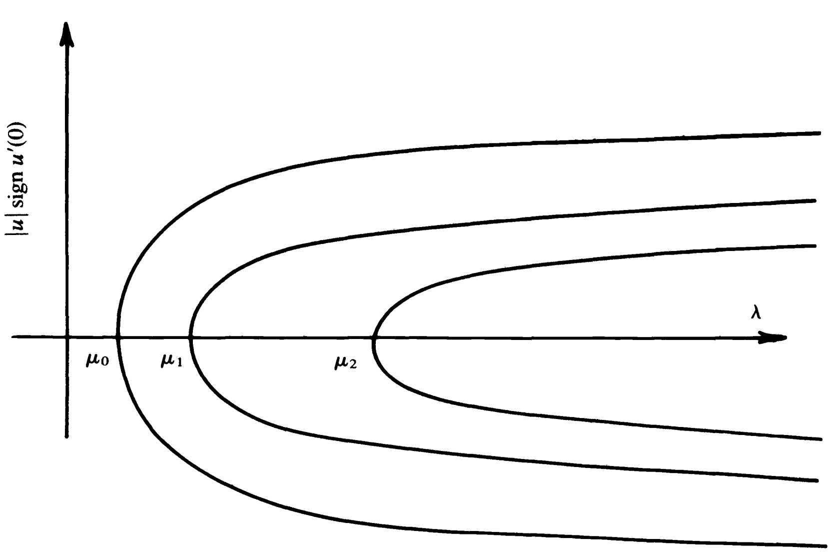

FIG. 1. *Bifurcation diagram for eq.* (1), *k odd.*

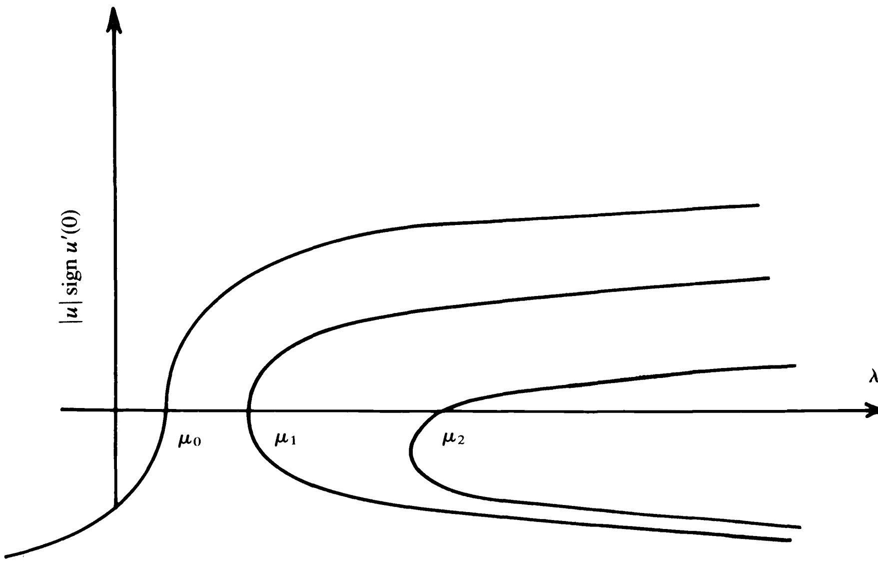

FIG. 2. *Bifurcation diagram for eq.* (1), *k even.*

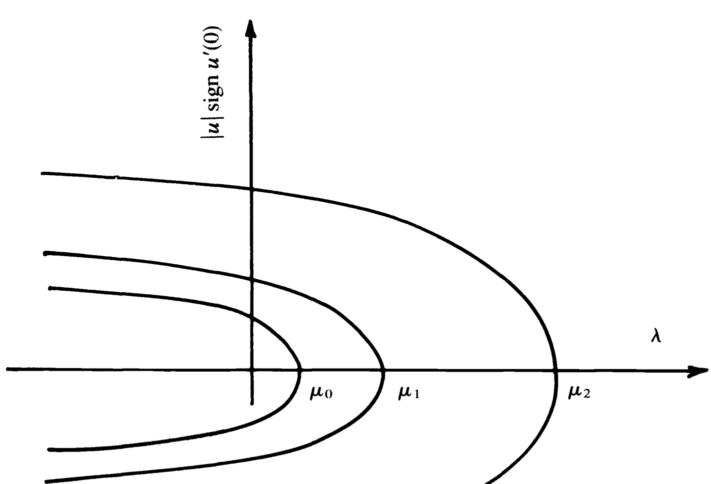

FIG. 3. *Bifurcation diagram for eq.* (1'), *k odd.*

than that of the subintervals where it is negative; moreover, it increases monotonically with λ *(for* $\lambda > \mu_{2m+1}$*). The same result holds true for solutions with* $2m$ *nodes and* $u'(0) > 0$ $(m \geqq 1)$.

3. Phase plane analysis. Let us rewrite equation (1) as a first order system:

$$u' = v \qquad (2)$$
$$v' = -\lambda u + u^k.$$

It is easily seen that the quantity

$$E = v^2 + \lambda u^2 - (2/(k+1))u^{k+1} \qquad (3)$$

is constant along the solutions of (2); for any fixed $\lambda \in \mathbb{R}$, equation (3) gives all the trajectories of the system in the phase plane (u, v) as the energy E varies in $\mathbb{R}$. Notice that such trajectories are symmetric with respect to the u-axis, and, if k is odd, with respect to the v-axis as well.

As for the critical points of (2), we have the following situation: when k is odd, if $\lambda < 0$, the only critical point $(0, 0)$ is a saddle; if $\lambda > 0$, $(0, 0)$ is a center and $(\pm\lambda^{1/(k-1)}, 0)$ are saddle points; on the other hand, when k is even, there are two critical points, $(0, 0)$ and $(\lambda^{1/(k-1)}, 0)$, for any $\lambda \in \mathbb{R}$: if $\lambda < 0$, $(0, 0)$ is a saddle and $(\lambda^{1/(k-1)}, 0)$ is a center; the reverse situation holds when $\lambda > 0$.

Due to the boundary conditions, we are actually interested just in trajectories which cross the v-axis at least twice (possibly at their endpoints), and the u-axis at least once; clearly, this requirement rules out the case k odd, $\lambda < 0$. In the remaining cases, a straightforward analysis gives, for any fixed λ, the trajectories depicted in Figs. 4–6.

An immediate consequence is that, when k is even and $\lambda < 0$, the only trajectories of the type required lie in the left half plane. In particular, in such case no positive solutions exist, nor solutions with nodes.

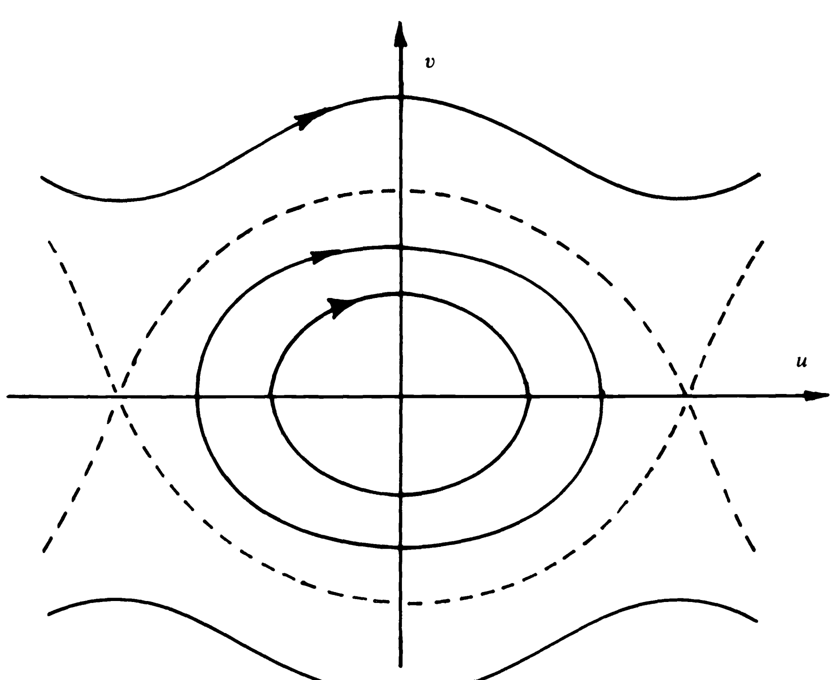

FIG. 4. *Phase plane trajectories for eq.* (1); *k odd,* $\lambda > 0$.

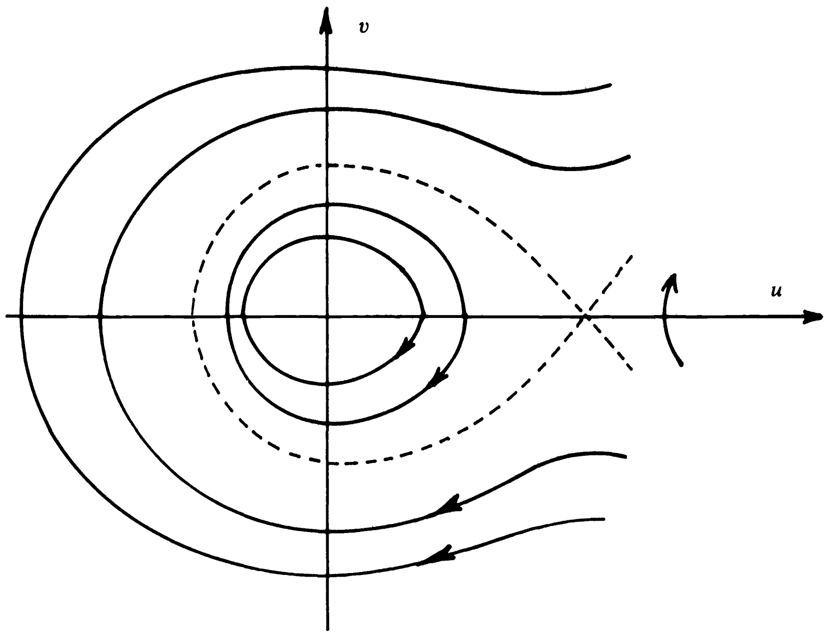

FIG. 5. *Phase plane trajectories for eq.* (1); *k even*, $\lambda > 0$.

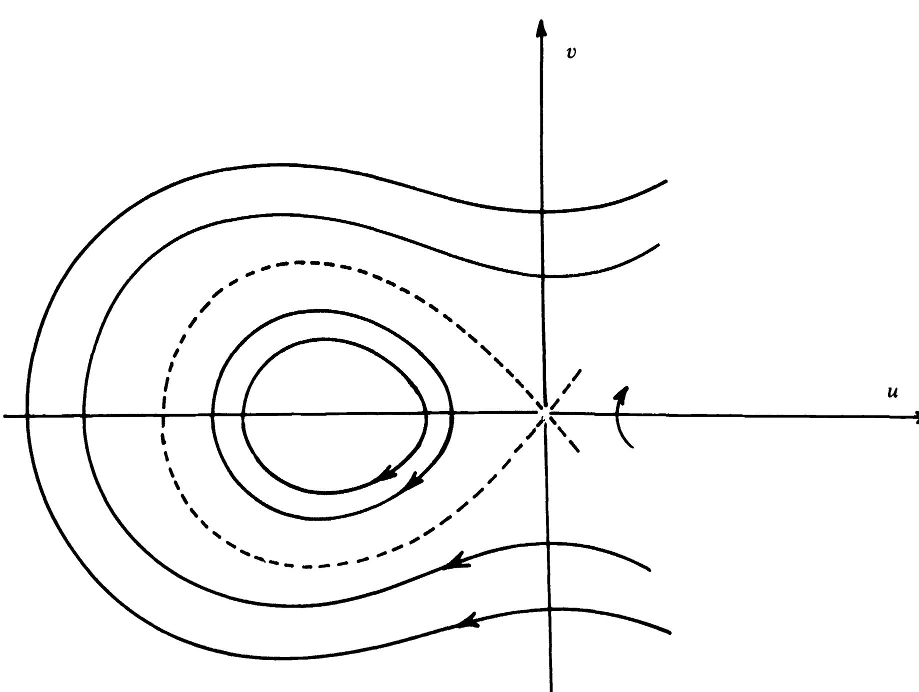

FIG. 6. *Phase plane trajectories for eq.* (1); *k even*, $\lambda < 0$.

The boundary conditions exclude as well: i) when k is odd, trajectories whose energy is larger than $E_* = ((k-1)/(k+1)) \cdot \lambda^{(k+1)/(k-1)}$ (that is, the energy of the degenerate orbits $(\pm\lambda^{1/(k-1)}, 0)$; ii) when k is even and $\lambda < 0$, trajectories with energy less than zero; iii) when k is even and $\lambda > 0$, trajectories with energy greater than E_*, lying in the right half plane.

A most important requirement must now be imposed on the remaining trajectories: in fact, in order that a trajectory corresponds to a solution of the problem (1), it must be covered in the prescribed time T. As we shall see, for any solution with a fixed number of nodes, this condition can be written as a (transcendental) equation for E; the occurrence of solutions for such equations will be successively characterized in terms of λ. It is again convenient to distrnguish two cases, according to the parity of k:

Case (a): k odd. Let us call $r_+(E, \lambda)$ the smallest positive root of the polynomial

$$u \to P(E, \lambda, u) = E - \lambda u^2 + (2/(k+1))u^{k+1}$$

(where $\lambda \in \mathbb{R}^+$): namely, for any λ, $r_+(E, \lambda)$ is the abscissa of the intersection of the orbit having energy E with the positive u axis. Then, if a solution of (1) having no nodes exists, the following relation must hold

$$\frac{T}{2} = \int_0^{r_+(E,\lambda)} \frac{du}{u'} = \int_0^{r_+(E,\lambda)} du(P(E, \lambda, u))^{-1/2}. \tag{4}$$

In fact, due to the u-axis symmetry, the integral at the right-hand side of (4) expresses half the time needed for a trajectory starting from $(0, \sqrt{E})$ to attain the point $(0, -\sqrt{E})$, having crossed the u-axis just once. Notice that the v-axis symmetry guarantees that this time is in fact the same both for positive and negative solutions. It is also easy to see that the condition (4) is not only necessary, but also sufficient for the existence of solutions without nodes: in fact, if a solution $\bar{E}$ of (4) exists, integrating the equation $u' = (P(\bar{E}, \lambda, u))^{1/2}$ provides a solution of this kind.

The characterization of solutions with m nodes ($m \geqq 1$) is obtained in a similar way: in fact, due to the u-, v-symmetries, a necessary and sufficient condition for the existence of a couple of solutions with m nodes is

$$\frac{T}{2} = (m+1) \int_0^{r_+(E,\lambda)} du(P(E, \lambda, u))^{-1/2} \tag{5}$$

as the corresponding trajectory crosses the u-axis $(m+1)$ times.

Case (b): k even. This case is more involved, essentially due to the lack of the v-axis symmetry. Let us define $r_+(E, \lambda)$, $r_-(E, \lambda)$ as the smallest (in absolute value) strictly positive, respectively negative root of the polynomial $P(E, \lambda, \cdot)$,[3] where $\lambda \in \mathbb{R}$.

A necessary and sufficient condition for the existence of a positive solution of (1) (that is, no nodes, and $u'(0) > 0$) is then

$$\frac{T}{2} = \int_0^{r_+(E,\lambda)} du\ (P(E, \lambda, u))^{-1/2}; \tag{6_+}$$

on the other hand, a negative solution of (1) exists if and only if

$$\frac{T}{2} = \int_{r_-(E,\lambda)}^{0} du\ (P(E, \lambda, u))^{-1/2}. \tag{6_-}$$

[3] In the Case (a), the role of r_- is played by $-r_+$.

As to the solutions with nodes, we shall distinguish between the cases of an even and of an odd number of nodes.

In fact, to a solution of (1) possessing $2m+1$ nodes, there corresponds a closed trajectory in the (u, v)-plane, which is covered $m+1$ times: as a consequence, such trajectories are covered in the same time, both if they are initiated at $(0, \sqrt{E})$, and at $(0, -\sqrt{E})$: hence the (necessary and sufficient) condition for the existence of solutions with $u'(0)>0$ and $u'(0)<0$ is the same, namely

$$\frac{T}{2}=(m+1)\int_{r_-(E,\lambda)}^{r_+(E,\lambda)} du\,(P(E,\lambda,u))^{-1/2}. \tag{7}$$

If there are $2m$ nodes, a further distinction, namely between solutions with $u'(0)>0$ and $u'(0)<0$, is needed. In the first case, a closed trajectory in the phase plane is covered m times, and its part lying in the right half plane is covered once more: thus we get the following necessary and sufficient condition

$$\frac{T}{2}=m\int_{r_-(E,\lambda)}^{r_+(E,\lambda)} du\,(P(E,\lambda,u))^{-1/2}+\int_0^{r_+(E,\lambda)} du\,(P(E,\lambda,u))^{-1/2}. \tag{8_+}$$

In the case $u'(0)<0$, on the other hand, the closed trajectory is covered again m times, but it is its part lying in the left half plane which is now covered once more; hence we arrive at the (necessary and sufficient) condition

$$\frac{T}{2}=m\int_{r_-(E,\lambda)}^{r_+(E,\lambda)} du\,(P(E,\lambda,u))^{-1/2}+\int_{r_-(E,\lambda)}^{0} du\,(P(E,\lambda,u))^{-1/2}. \tag{8_-}$$

Proving Theorems 1 and 2 amounts to finding necessary and sufficient conditions on λ in order that the equations for E (4), (5), ($6_\pm$), (7), ($8_\pm$) have solutions. This will be done in the following section, by establishing monotonicity properties of the right hand sides of these equations.

4. Proof of the results. Let us prove Theorem 2.[4] To this end, we shall present a lemma concerning some useful properties of the roots r_+, r_-. Notice that r_+, as a function of (E, λ), is defined only on $Q_+=\{(E,\lambda)\in(0,E_*)\times\mathbb{R}^+,\ E_*=((k-1)\div(k+1))\lambda^{(k+1)/(k-1)}\}$; as for r_-, we shall be dealing only with its restriction to $Q_-=\mathbb{R}^+\times\mathbb{R}$.

LEMMA 1. *The following properties are valid*:

(i) $0<r_+(E,\lambda)<\lambda^{(1/(k-1)}$ *for all* $(E,\lambda)\in Q_+$;

(ii) $r_-(E,\lambda)<0$ *for all* $E\in\mathbb{R}^+$ *and* $\lambda\in\mathbb{R}^-$, *and*
$r_-(E,\lambda)<\lambda^{1/(k-1)}<0$ *for all* $E\in\mathbb{R}^+$ *and* $\lambda\in\mathbb{R}^-$;

(iii) $\dfrac{\partial}{\partial E}r_+(E,\lambda)>0$ *for all* $(E,\lambda)\in Q_+$;

(iv) $\dfrac{\partial}{\partial E}r_-(E,\lambda)<0$ *for all* $(E,\lambda)\in Q_-$.

Proof. Items (i), (ii) follow from an elementary analysis of the polynomial $P(E,\lambda,\cdot)$; as for (iii), (iv), suffice it to make use of (i), (ii) and of the relation

$$\frac{\partial}{\partial E}r_\pm(E,\lambda)=\{2r_\pm(E,\lambda)(\lambda-r_\pm^{k-1}(E,\lambda))\}^{-1},$$

[4] The proof of Theorem 1 will be omitted, as the arguments needed are of the same type—and much simpler: in fact, due to the v-symmetry, only the integral $\int_0^{r_+(E,\lambda)} du(P(E,\lambda,u))^{-1/2}$ is involved.

which is obtained by differentiating with respect to E the identity $P(E, \lambda, r_\pm(E, \lambda)) = 0$.

Let us now consider the following functions

$$I_\pm:\ Q_\pm \to \mathbb{R}^+,\quad (E, \lambda) \to I_\pm(E, \lambda) \doteq \pm \int_0^{\pm r_\pm(E,\lambda)} du\ ((P(E, \lambda, u))^{-1/2};$$

$$I:\quad Q_+ \to \mathbb{R}^+,\quad (E, \lambda) \to I(E, \lambda) \doteq I_+(E, \lambda) + I_-(E, \lambda);$$

$$Z_m: Q_+ \to \mathbb{R}^+,\quad (E, \lambda) \to Z_m(E, \lambda) \doteq mI(E, \lambda) + I_-(E, \lambda) \qquad (m \in \mathbb{N}).$$

Notice that I_- (resp. I_+, I, Z_m) can be defined for all $\lambda \in \mathbb{R}$ (resp., $\lambda \in \mathbb{R}^+$).

LEMMA 2.

(i) $I_+(\cdot, \lambda)$ *is strictly increasing on* $(0, E_*)$, *for all* $\lambda \in \mathbb{R}^+$;

(ii) $I_-(\cdot, \lambda)$ *is strictly decreasing on* $\mathbb{R}^+$, *for all* $\lambda \in \mathbb{R}$;

(iii) $I(\cdot, \lambda)$ *is strictly increasing on* $(0, E_*)$, *for all* $\lambda \in \mathbb{R}^+$;

(iv) $\lim_{E\to 0^+} I_+(E, \lambda) = \pi/(2\sqrt{\lambda})$; $\lim_{E\to E_*^-} I_+(E, \lambda) = +\infty$ *for all* $\lambda \in \mathbb{R}^+$;

(v) $\lim_{E\to 0^+} I_-(E, \lambda) = \pi/(2\sqrt{\lambda})$ (*for all* $\lambda \in \mathbb{R}^+$), *and* $\lim_{E\to +\infty} I_-(E, \lambda) = 0$ *for all* $\lambda \in \mathbb{R}$.

Proof. As for (i), (ii), observe that, subtracting in the expression under square root in the definition of $I_\pm$ the quantity $P(E, \lambda, r_\pm) = 0$ and performing the change of integration variable $u = r_+ t$, we arrive at the following expression

$$I_\pm(E, \lambda) = \int_0^1 dt\{(1-t)[\lambda(1+t) - (2/(k+1))r_\pm^{k-1}(E, \lambda)(1+t+\cdots t^k)]\}^{-1/2}$$

whence the monotonicity properties of I_+, I_- follow from Lemma 1. As for (iii), we shall limit ourselves, for the sake of simplicity, to the case $k = 2$. Denote by $r_3(E, \lambda)$ the third root of the polynomial $P(E, \lambda, \cdot)$, which is real for $E \in (0, E_*)$ and larger than λ ($\lambda \in \mathbb{R}^+$); using Cardano's formulae we obtain the explicit representation:

$$r_+(E, \lambda) = \tfrac{1}{2}\lambda + \lambda \cos((\varphi + 4\pi)/3),$$
$$r_-(E, \lambda) = \tfrac{1}{2}\lambda + \lambda \cos((\varphi + 2\pi)/3),$$
$$r_3(E, \lambda) = \tfrac{1}{2}\lambda + \lambda \cos(\varphi/3),$$

where $\varphi = \arccos((E_* - 2E)/E_*)$ varies between 0 and π as E varies between 0 and E_*. On the other hand, putting the complete elliptic integral $I(E, \lambda)$ in the Legendre canonical form and differentiating with respect to E, we get

$$\frac{\partial}{\partial E} I(E, \lambda) = ((E_* - E)E)^{-1/2}(\sqrt{3}/2) \cdot \int_0^1 dt\Big\{(1-t^2)^{-1/2}[(r_3 - r_-) - (r_+ - r_-)t^2]^{-3/2} \cdot \frac{\partial}{\partial \varphi}[(r_+ - r_-)t^2 - (r_3 - r_-)]\Big\}(E, \lambda).$$

Observe now that

$$\frac{\partial}{\partial \varphi}(r_+ - r_-) = (\lambda/\sqrt{3}) \cos(\varphi/3) > 0, \quad \forall \varphi \in (0, \pi),$$

and

$$\frac{\partial}{\partial \varphi}(r_3 - r_-) = -(\lambda/\sqrt{3}) \cos((\varphi + 4\pi)/3) < 0, \quad \forall \varphi \in (\pi/2, \pi),$$

so that $I(\cdot,\lambda)$ is strictly increasing on the subinterval $(E_*/2, E_*)$. As to the subinterval $(0, E_*/2)$, use is made of the estimate

$$\frac{\partial}{\partial E} I(E,\lambda) \geqq \left(\frac{\pi}{4}\right)\left[\left(r_3-\frac{\lambda}{2}\right)(r_3-r_-)^{-3/2}+(2r_+-\lambda)\cdot(r_3-r_+)^{-3/2}\right](E,\lambda).$$

In fact, it is easy to show that the above lower bound is strictly positive for $E \in (0, E_*/2)$, which completes the proof of (iii). The proofs of (iv) and (v) are straightforward and will be omitted.

LEMMA 3. *The function $E \to Z_m(E,\lambda)$ is decreasing in a right neighborhood of zero; moreover, there exists a positive decreasing sequence $\{\varepsilon_m\}$ such that $Z_m(\cdot,\lambda)$ is increasing for $E > \varepsilon_m$, and $\lim_{E\to E^*_-} Z_m(E,\lambda) = +\infty$ ($m \in \mathbb{N}$, $\lambda \in \mathbb{R}^+$).*

The proof follows from a lengthy calculation, similar to that of Lemma 2 part (iii), and the recurrence relation $Z_m = Z_{m-1} + I$. Pursuing the analytical investigation of further properties of Z_m is extremely involved; thus numerical computations[5] were performed to establish the following result:

For any $m \in \mathbb{N}$, $\lambda \in \mathbb{R}^+$, $Z_m(\cdot,\lambda)$ possesses a unique minimum point ζ_m, where $\{\zeta_m\}$ is a monotonic sequence, decreasing to zero, and $\{Z_m(\zeta_m,\lambda)\}$ is increasing.

Proof of Theorem 2. We just sketch the proof of (i), the other cases being similar. According to Lemma 2, $I_+(\cdot,\lambda)$ is strictly increasing, and $\lim_{E\to 0^+} I_+(E,\lambda) = \pi/(2\sqrt{\lambda})$, so that a unique solution of (6_+) exists if and only if $T/2 > \pi/(2\sqrt{\lambda})$, i.e., if and only if $\lambda > \mu_0$.

Proof of Proposition 1. Let us first investigate the case k odd: the supremum norm of the nontrivial solution with m nodes is precisely $r_+(E,\lambda)$, where, for any λ, E is determined by the equation (5) above. Using the same procedure as in the proof of Lemma 2, we may rewrite (5) as follows:

$$\text{(9)}\quad \begin{aligned} 0 = \lambda^{-1/2}\int_0^1 dt\,\{(1-t)[(1+t)-(2r_+^{k-1}/(k+1)\lambda)\\ \cdot(1+t+\cdots+t^k)]\}^{-1/2} - T/(2(m+1)) \qquad (\lambda > \mu_m).\end{aligned}$$

It is easily seen that (9) defines r_+ as an implicit function of λ, whose derivative is

$$\text{(10)}\quad \frac{d}{d\lambda}r_+ = \frac{k+1}{k-1}\frac{r_+^{k-2}}{2} \cdot \frac{\displaystyle\int_0^1 dt\,(1-t)^{-1/2}(1+t)\left[(1+t)-\frac{2}{k+1}\frac{r_+^{k-1}}{\lambda}(1+t+\cdots+t^k)\right]^{-3/2}}{\displaystyle\int_0^1 dt\,(1-t)^{-1/2}(1+t+\cdots+t^k)\left[(1+t)-\frac{2}{k+1}\frac{r_+^{k-1}}{\lambda}(1+t+\cdots+t^k)\right]^{-3/2}};$$

hence r_+ increases with λ. According to Lemma 1, as λ goes to infinity, r_+ cannot diverge faster than $\lambda^{1/(k-1)}$; in fact, it diverges exactly as $\lambda^{1/(k-1)}$: namely, otherwise it would follow from (9) that $T/(2(m+1)) \leqq (\pi/2)\lim_{\lambda\to+\infty}\lambda^{-1/2} = 0$, which is absurd.

---

[5] It was found convenient to choose as independent variable the following (strictly increasing) function of λ:

$$\theta = \arcsin\{2(R^2-1)^{1/2}/(\sqrt{3}+(R^2-1)^{1/2})\}^{1/2}, \quad \text{where } R = 2\lambda/(\lambda-2r_3).$$

The computations, performed on a computer PDP 11/40, provided evidence of the following facts: (i) I is a convex function of θ; (ii) there is a $m' \in \mathbb{N}$ such that $Z_{m'}$ is a convex function of θ. Due to the recurrence formula for Z_m it follows that, for $m \geqq m'$, Z_m has at most one—and therefore exactly one—minimum. As to the Z_m's with $m < m'$, a direct calculation shows that they have a unique minimum as well.

Let us now turn to the case k even. For solutions without nodes, the procedure is the same as in the case k odd; for solutions with an odd number of nodes, or with an even number of nodes and $u'(0) > 0$, consider the following functions:

$$Y_m(E, \lambda) = \begin{cases} \frac{1}{2}(m+1)I(E, \lambda) - \frac{1}{2}T & \text{if } m \text{ is odd;} \\ mI(E, \lambda) + I_+(E, \lambda) - \frac{1}{2}T & \text{if } m \text{ is even.} \end{cases}$$

It is possible to think of E as a function of λ, $E(\lambda)$, implicitly defined through the equation $Y_m(E, \lambda) = 0$; now, the maximum norm of the solutions under consideration is $r_+(E(\lambda), \lambda)$ (in fact $r_+ \geqq |r_-|$ on Q_+), and a direct calculation shows that $(d/d\lambda)r_+(E(\lambda), \lambda) > 0$. The claims on the asymptotic behavior follow as in the case k odd.

Proof of Proposition 2. It is an easy consequence of (10) in the case k odd and in the case k even, solutions without nodes. The remaining cases are similar.

Proof of Proposition 3. According to Lemma 2, $I_+(E, \lambda) > I_-(E, \lambda)$ on Q_+. On the other hand, the time needed to cover the part of the trajectory lying in the right (resp. left) half plane is $2I_+$ (resp. $2I_-$), whence the first claim is proven. To complete the proof, suffice it to compute $(d/d\lambda)I_+(E(\lambda), \lambda)$, which turns out to be positive.

Acknowledgment. We are indebted to the referee, who suggested several improvements of the original manuscript.

REFERENCES

[1] P. H. RABINOWITZ, *Nonlinear Sturm–Liouville problems for second order differential equations*, Comm. Pure Appl. Math., **23** (1970), pp. 939–961.

[2] ———, *Some aspects of nonlinear eigenvalue problems*, Rocky Mountain. J. Math., 2 (1973), pp. 161–202.

[3] P. DE MOTTONI AND A. TESEI, *On a class of nonlinear eigenvalue problems*, Boll. Un. Mat. Ital., 14 B, (1977), pp. 172–189.

SIAM J. MATH. ANAL.
Vol. 9, No. 6, December 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0004 $01.00/0

ON THE ASYMPTOTIC SOLUTION OF A PARTIAL DIFFERENTIAL EQUATION WITH AN EXPONENTIAL NONLINEARITY*

V. H. WESTON†

Abstract. The asymptotic behavior of a large norm (maximum) solution of the Dirichlet problem associated with the equation

$$-\Delta u = \lambda e^u$$

for a bounded simply-connected domain in $\mathbb{R}^2$ is investigated for the case of the positive parameter λ tending to zero. By means of a conformal transformation function $f(z)$, the problem is transformed to one involving the unit disc. For a class of domains which are described by implicit conditions for $f(z)$, a first and higher asymptotic expressions are developed for the large norm solution characterized by a single maximum proportional to $\ln(1/\lambda)$. It is shown that for λ sufficiently small, an exact solution can be generated by the modified Newton iteration scheme, if the asymptotic solution of appropriate order is used for the initial step.

1. Introduction. We will consider the nonlinear problem (P) in $\mathbb{R}^2$

$$\text{(P)} \qquad \begin{aligned} -\Delta u &= \lambda e^u, \qquad y \in D \\ u &= 0, \qquad y \in \delta D \end{aligned}$$

where $y = (y_1, y_2)$, Δ is the Laplacian operator, D is a compact simply-connected domain with smooth boundary δD, and the parameter λ is such

$$\lambda > 0.$$

Some general results on the solutions of the system (P) have been given in the literature [1], [2], [3], [4]. There exist multiple solutions for $0 < \lambda < \lambda^*$, where λ^* is a bifurcation point. There is a low norm or "minimal" solution which tends to zero as $\lambda \to 0$. All solutions are pointwise nonnegative. For $\lambda > \lambda^*$ there is no solution.

It is the purpose of this paper to obtain an asymptotic approximation for the large norm [1] solution as $\lambda \to 0$, with emphasis on the solution that has a single maximum. This will be achieved by utilizing the Liouville form of the solution of the differential equation given by [5], [6].

$$\lambda e^u = \frac{8|F'(w)|^2}{[1 + |F(w)|^2]^2} \tag{1.1}$$

where $F(w)$ is a function of $w = y_1 + iy_2$. This result is obtained by transforming the partial differential to Liouville's form

$$u_{\xi\eta} = -\frac{\lambda}{4} e^u,$$

where

$$\xi = y_1 + iy_2, \qquad \eta = y_1 - iy_2.$$

Upon differentiating the Liouville form of the equation with respect to the variable ξ, and then eliminating the exponential term, one obtains a differential equation which

* Received by the editors May 24, 1976, and in final revised form March 28, 1977.
† Division of Mathematical Sciences, Purdue University, West Lafayette, Indiana 47907.

can be directly integrated to yield

$$u_{\xi\xi}-\tfrac{1}{2}u_\xi^2=B_1(\xi),$$

where $B_1(\xi)$ is an analytic function. Similarly one obtains

$$u_{\eta\eta}-\tfrac{1}{2}u_\eta^2=B_2(\eta),$$

where $B_2(\eta)$ is an anti-analytic function.

Upon integration one obtains the general solution

$$e^u=\frac{8}{\lambda}\frac{\varphi'(\xi)\psi'(\eta)}{[\varphi(\xi)\psi(\eta)+1]^2}.$$

With

$$\varphi(\xi)=F(y_1+iy_2), \qquad \psi(\eta)=\overline{F(y_1+iy_2)},$$

the right-hand side of the above equation is real and positive, and yields results (1.1).

The restriction upon the function F is found by noting that

$$B_1(\xi)=\left(\frac{F''}{F'}\right)'-\frac{1}{2}\left(\frac{F''}{F'}\right)^2$$

must be an analytic function. It can be shown from this that for the domain under consideration, F must have the following properties

(i) apart from simple poles F is analytic,

(ii) the zeros of F must be simple,

(iii) F' must have no zeros in the domain.

Thus it is seen that F is a meromorphic function with simple zeros and poles.

To simplify analysis it will be convenient to transform the domain D to a unit disc by the use of a conformal transformation.

DEFINITION. Let $w=f(z)$, $z=x_1+ix_2$ be the conformal mapping that transforms the interior of D into the unit disc $|z|<1$.

Problem (P) now reduces to the following problem (P′)

$$\text{(P}'\text{)}\qquad \begin{aligned} -\Delta u&=\lambda|f'(z)|^2 e^u, && |z|<1,\\ u&=0, && |z|=1.\end{aligned}$$

In terms of the Liouville representation, the solution of the differential equation associated with problem (P′) is given by

$$\lambda e^u=\frac{8|f'(z)|^{-2}|G'(z)|^2}{[1+|G(z)|^2]^2} \tag{1.2}$$

where $G(z)=F[f(z)]$. The meromorphic function $G(z)$ is chosen so that

$$\lambda|f'(z)|^2=\frac{8|G'(z)|^2}{[1+|G(z)|^2]^2}, \quad \text{for } |z|=1. \tag{1.3}$$

For the special case where D is the unit disc, hence $f'(z)\equiv 1$, the solutions are given by

$$G(z)=cz \quad \text{or} \quad cz^{-1},$$

where c is a constant, determined by the boundary condition. The large norm solution is given explicitly by

$$G(z)=\beta^{1/2}z \tag{1.4}$$

with

$$\beta=\frac{4}{\lambda}\left[1-\frac{\lambda}{4}+\left(1-\frac{\lambda}{2}\right)^{1/2}\right]. \tag{1.5}$$

Note that as $\lambda\to 0$, $\beta\to 8/\lambda$ and

$$G(z)\sim\left(\frac{8}{\lambda}\right)^{1/2} z \quad \text{as } \lambda\to 0.$$

Thus the large norm solution has the property that $|G(z)|\gg 1$ everywhere in the unit disc except in the neighborhood of $z=0$. The generalization of this result will be the basis for the construction of the large norm solution for a general class of domains D. In particular we will consider those domains D such that their transformation function $f(z)$ which maps D into the unit disc $|z|<1$, has the property given below.

We will assume that there exists a δ, $|\delta|<1$, such that

$$\bar{\delta}=\tfrac{1}{2}(1-|\delta|^2)\frac{f''(\delta)}{f'(\delta)}. \tag{1.6}$$

The general question of existence and uniqueness of the solution to Equation (1.6), and the corresponding conditions on $f'(z)$, will not be considered here, and remains to be pursued. However we need to demonstrate that the class of functions $f'(z)$ for which δ exists is not vacuous.

If $f''(z)$ has a zero at the origin, then we can take $\delta=0$; otherwise rewrite Equation (1.6) in the form

$$\frac{\delta f''(\delta)}{f'(\delta)}=\frac{2|\delta|^2}{(1-|\delta|^2)}. \tag{1.7}$$

The problem then reduces to finding a nonzero δ, lying on the curve

$$\operatorname{Im}\frac{zf''(z)}{f'(z)}=0, \qquad |z|<1, \tag{1.8}$$

such that its real part is given by the right-hand side of (1.7). If $f''(z)$ has no zeros in $|z|<1$, then it can be shown that the curve (1.8) goes through the origin and intersects the unit circle $|z|=1$ in two points. Since the right-hand side of (1.7) varies between zero and ∞ as $|\delta|$ varies between zero and one, it is obvious that Equation (1.7) will have a solution. The solution is unique if the curve (1.8) is not tangent to the circles $|z|=\text{constant}$.

If more than one δ exists, we will not specify, initially, which one to take. (Note that there exists the possibility that multiple values of δ will lead to solutions with multiple maxima. Since in this analysis we are interested in solutions with a single large maximum, we will leave this question to future effort.)

In § 2 we will obtain the first order, then the higher order asymptotic approximations to the solution of problem (P′) that has a single maximum. In obtaining the higher approximations, additional implicit constraints are placed upon the transformation function $f'(z)$, and choice of δ, if more than one exists. In § 3, we will consider the conditions for which a Newton-type iteration scheme can be used to generate the exact solution. We transform the problem (P′) to an integral form, and in the iteration process use the previously developed asymptotic approximations as the initial approximation. It is shown that as $\lambda\to 0$, the first order approximation is not good enough for the iteration process to converge.

2. Construction of asymptotic approximation.

2.1. First order solution. To obtain a first order asymptotic approximation to the large norm solution for problem (P′) for $\lambda \to 0$, we will use Liouville's form given by (1.2) and set

$$G(z,\lambda)=\lambda^{-1/2}G_0(z),$$

where $G_0(z)$ is independent of λ. Equation (1.3) for the boundary condition reduces to

$$|f'(z)|^2=8\frac{|G_0'(z)|^2}{|G_0(z)|^4}[1+O(\lambda)]$$

for $|z|=1$. We shall neglect the term $O(\lambda)$, in which case we have

$$\sqrt{8}\frac{d}{dz}[G_0(z)]^{-1}=A(z)f'(z), \qquad |z|=1, \tag{2.1}$$

where $|A(z)|=1$ for $|z|=1$.

We shall choose $A(z)$ so that Equation (2.1) holds in the closed disc $|z|\leqq 1$, and such that apart from possible simple poles, $G_0(z)$ is analytic in the disc and has one zero. Set

$$A(z)=\left(\frac{1-\bar{\delta}z}{z-\delta}\right)^2. \tag{2.2}$$

Because of assumption on δ given by (1.6), it follows that Equation (2.1), now valid for the unit disc, can be integrated to yield

$$\sqrt{8}[G_0(z)]^{-1}=P_0(z)/(z-\delta)+C_0, \tag{2.3}$$

where

$$\begin{aligned}P_0(z)=-(1-|\delta|^2)^2f'(\delta)+(z-\delta)Q(z)-(z-\delta)^2 3(\bar{\delta})^2 f'(\delta)\\ +\frac{(z-\delta)^3 f'(\delta)(\bar{\delta})^3}{(1-|\delta|^2)}\end{aligned} \tag{2.4}$$

with

$$Q(z)=\int_\delta^z [f'(\xi)-f'(\delta)-(\xi-\delta)f''(\delta)]A(\xi)\,d\xi$$

and C_0 is an arbitrary constant. For future use we need the following quantities

$$\begin{aligned}P_0(\delta)&=-(1-|\delta|^2)^2 f'(\delta),\\ P_0''(\delta)&=-6(\bar{\delta})^2 f'(\delta)+f'''(\delta)(1-|\delta|^2)^2.\end{aligned} \tag{2.5}$$

The constant C_0 will be uniquely determined when we come to construct higher order solutions. However for the first order solution given below we will take C_0 to be zero. Hence we take u_1 to be the first order solution given by

$$e^{u_1}=\frac{8|f'(z)|^{-2}|G_0'(z)|^2}{[\lambda+|G_0(z)|^2]^2},$$

which simplifies to

$$e^{u_1}=|1-z\bar{\delta}|^4\left[|z-\delta|^2+\frac{\lambda}{8}|P_0(z)|^2\right]^{-2}. \tag{2.6}$$

It is interesting to note that at $z=\delta$, this solution has the value

$$u_1(\delta)=-2\ln\left[\frac{\lambda|f'(\delta)|^2}{8}\right].$$

The results can be summarized as follows

THEOREM. *The first order solution u_1 given by* (2.6) *satisfies the equation*

$$-\Delta u_1=\lambda|f'(z)|^2 e^{u_1}, \qquad |z|<1, \tag{2.7}$$

and the boundary condition

$$u_1\sim O(\lambda), \qquad |z|=1,$$

asymptotically as $\lambda\to 0$. Furthermore,

$$\max_{|z|\leqq 1}|u_1|\sim O\left(\ln\left(\frac{1}{\lambda}\right)\right), \qquad \lambda\to 0.$$

2.2. Second and higher order solutions. To obtain an explicit expression for the higher order approximation, we will set

$$G(z)=\lambda^{-1/2}\sum_{n=0}\lambda^n G_n(z)$$

in the Liouville expression where $G_n(z)$ are independent of λ. As in the previous section $G_0(z)$ will satisfy the equation

$$\sqrt{8}|G_0'(z)|=|G_0(z)|^2|f'(z)|, \qquad |z|=1, \tag{2.8}$$

and its solution will be given by Equation (2.3) with the constant C_0 to be specified. For the first order approximation we took C_0 to be zero, but for the higher order approximation it will have to be chosen in a specific way.

The boundary condition (1.3) with $|z|=1$, becomes

$$|f'(z)|^2\left[\lambda+\left|\sum_{n=0}\lambda^n G_n(z)\right|^2\right]^2=8\left|\sum_{n=0}\lambda_n G_n'(z)\right|^2. \tag{2.9}$$

For simplification of analysis set

$$B_n=G_n(z)/G_0(z), \qquad B_n'=G_n'(z)/G_0'(z),$$

and

$$b_n=\sum_{m=0}^{n}B_m\bar{B}_{n-m}, \qquad b_n'=\sum_{m=0}^{n}B_m'\bar{B}_{n-m}'$$

Then we have

$$\left|\sum_{n=0}\lambda^n G_n(z)\right|^2=|G_0(z)|^2\sum_{n=0}\lambda^n b_n.$$

By equating coefficients of λ^n in (2.9) and using (2.8) we obtain for $|z|=1$,

$$\sum_{p=0}^{n}\left(b_p+\frac{1}{|G_0|^2}\delta_{1p}\right)\left(b_{n-p}+\frac{1}{|G_0|^2}\delta_{1n-p}\right)=b_n' \tag{2.10}$$

where δ_{1p} is the Kronecker delta. Expressed in terms of G_n, (2.10) reduces to the general form

$$\text{Real}\left(\frac{G_n'}{G_0'}-2\frac{G_n}{G_0}\right)=h_n(z), \qquad |z|=1, \tag{2.11}$$

where

$$h_1(z)=1/|G_0(z)|^2, \tag{2.12}$$

$$h_2(z)=\left|\frac{G_1}{G_0}\right|^2-\frac{1}{2}\left[I_m\frac{G_1'}{G_0'}\right]^2, \tag{2.13}$$

and for $n \geqq 3$,

$$h_n(z)=\text{Real}\left[-\frac{G_{n-1}'\bar{G}_1'}{|G_0'|^2}+2\frac{G_{n-1}\bar{G}_1}{|G_0|^2}\right]+2\text{ Real}\left(\frac{G_{n-1}}{G_0}\right)\text{Real}\left(\frac{G_1'}{G_0'}\right)+m_n(z) \tag{2.14}$$

where $m_n(z)$ is a function of lower order terms $G_{n-2}, G_{n-3}, \cdots$ etc. Its explicit form will not be given here.

The problem reduces to finding $G_n(z)$, $n=1,2,3\cdots$ satisfying (2.11), such that G_n/G_0 and G_n'/G_0' are analytic in the unit disc, and to determine the constant C_0 associated with $G_0(z)$.

Let $H_n(z)$ be the function which is analytic on the unit disc, and such that its real part on $|z|=1$ is given by $h_n(z)$. Apart from an arbitrary imaginary constant, $H_n(z)$ is uniquely determined. It will not be shown here, but such arbitrary constants combine to give an exponential factor with purely imaginary exponent in the expression for $G(z)$, and hence can be ignored. Thus $H_n(z)$ is given by

$$H_n(z)=\frac{-1}{\pi}\int_0^{2\pi}\frac{h_n(e^{i\theta'})}{(z-e^{i\theta'})}e^{i\theta'}\,d\theta'-\frac{1}{2\pi}\int_0^{2\pi}h_n(e^{i\theta'})\,d\theta'.$$

Equation (2.11) now can be written in the form

$$\frac{G_n'}{G_0'}-2\frac{G_n}{G_0}=H_n(z), \qquad |z|\leqq 1, \tag{2.15}$$

which can be solved to yield

$$\frac{G_n}{G_o}=-H_n(z)+G_0(z)\int_\delta^z\frac{H_n'(\xi)}{G_0(\xi)}d\xi+C_nG_0(z), \tag{2.16}$$

where C_n is an arbitrary constant. Since $G_0(\xi)$ has a simple zero at $\xi=\delta$, and we want an analytic solution, we shall place the following restriction on H_n,

$$H_n'(\delta)=0. \tag{2.17}$$

It will be shown that this yields a condition for determining the unknown constant C_{n-1}. We will consider the case $n=1$ first.

From (2.3) and (2.12) we have

$$h_1(z)=\frac{1}{8}\left|\frac{P_0(z)}{(z-\delta)}+C_0\right|^2, \qquad |z|=1.$$

Using the following result valid for $|z|=1$,

$$\text{Real}\left[\frac{P_0(\delta)\overline{C_0}}{(z-\delta)}\right]=\text{Real}\left[\frac{zC_0\bar{P}_0(\delta)}{(1-z\bar{\delta})}\right]$$

one can place the above result in the form

$$h_1(z)=\text{Real}\left\{\frac{|C_0|^2}{8}+\frac{1}{4}\frac{[P_0(z)-P_0(\delta)]}{(z-\delta)}\bar{C}_0+\frac{1}{4}\frac{zC_0\bar{P}_0(\delta)}{(1-z\bar{\delta})}\right\}+\frac{1}{8}\left|\frac{P_0(z)}{(z-\delta)}\right|^2.$$

It follows that

$$H_1(z)=\frac{|C_0|^2}{8}+\frac{1}{4}\frac{[P_0(z)-P_0(\delta)]}{(z-\delta)}\bar{C}_0+\frac{1}{4}\frac{zC_0\bar{P}_0(\delta)}{(1-z\bar{\delta})}$$
$$(2.18)\qquad -\frac{1}{8\pi}\int_0^{2\pi}\frac{|P_0(e^{i\theta'})|^2}{|e^{i\theta'}-\delta|^2}\left\{\frac{e^{i\theta'}}{z-e^{i\theta'}}+\frac{1}{2}\right\}d\theta'.$$

Thus (2.17) reduces to the form for $n=1$

$$-\bar{f}'(\delta)C_0+\tfrac{1}{2}P_0''(\delta)\bar{C}_0=Z(\delta),$$

where

$$(2.19)\qquad Z(\delta)=\frac{i}{2\pi}\oint\frac{|P_0(z')|^2}{|z'-\delta|^2(z'-\delta)^2}\,dz',$$

with the contour being the unit circle. Provided that the determinant $D(\delta)$ is not zero, where

$$(2.20)\qquad D(\delta)=|f'(\delta)|^2-\tfrac{1}{4}|f'''(\delta)(1-|\delta|^2)^2-6(\bar{\delta})^2f'(\delta)|^2,$$

the above equation can be solved for C_0 giving

$$(2.21)\qquad C_0=\frac{-1}{D}[f'(\delta)Z+\tfrac{1}{2}P_0''(\delta)\bar{Z}].$$

With C_0 specified by (2.21), $H_1(z)$ has the property that $H_1'(\delta)=0$, and the integral in the following for $G_1(z)$

$$(2.22)\qquad \frac{G_1(z)}{G_0(z)}=-H_1(z)+G_0(z)\left\{\int_\delta^z\frac{H_1'(\xi)}{G_0(\xi)}\,d\xi+C_1\right\}$$

describes an analytic function. However we require that G_1/G_0 and G_0'/G_0' be analytic in the unit disc. In this connection, a problem occurs if

$$(2.23)\qquad P(z)=P_0(z)+(z-\delta)C_0,$$

which is the denominator in the expression for $G_0(z)$, has a zero there. However, if $P(z)$ has only one zero, then the constant C_1 can be chosen so that the term in the square brackets in expression (2.22) vanishes at this zero. If $P(z)$ has no zero, then C_1 remains an arbitrary constant independent of λ, at least as far as the second order solution is concerned.

For the two cases where the analytic function $P(z)$ has no zero, or just one zero on the closed unit disc, we can define the second order solution u_2 as follows:

$$(2.24)\qquad \exp u_2=\frac{|1-z\bar{\delta}|^4|1+2\lambda(G_1/G_0)+\lambda H_1|^2}{\{(\lambda/8)|P(z)|^2+|z-\delta|^2|1+\lambda(G_1/G_0)|^2\}^2},$$

where G_1/G_0, H_1, $P(z)$ are given by (2.22), (2.18) and (2.23) respectively. The constant C_1 in expression (2.23) is arbitrary except when $P(z)$ has a single zero, in which case it is chosen to make G_1/G_0 analytic.

We will summarize these intermediate results in the following manner.

THEOREM. *For those transformations $f'(z)$, restricted by the implicit conditions*
(i) $D(\delta)\neq 0$,
(ii) *$P(z)$ has at most, one zero in the closed unit disc,*
the second order solution u_2 given by (2.24) *satisfies the relations*

$$-\Delta u_2=\lambda|f'(z)|^2 e^{u_2}, \qquad |z|<1,$$
$$u_2=O(\lambda^2), \qquad |z|=1.$$

2.3. *n*th order solution ($n \geqq 3$). We shall show how the previous construction can be used to obtain the higher order solutions, where the following implicit restrictions are placed on $f'(z)$,
(i) $D(\delta)\neq 0$,
(ii) $P(z)$ has no zeros in the unit disc.
The second condition implies that $G_0(z)$ is analytic in the unit disc.

The construction is the same as previously indicated except that the constant C_1 is chosen so that $H_2'(\delta)=0$, and in a similar manner the remaining constants C_{n-1} are chosen so that $H_n'(\delta)=0$. All that is needed is to show that no new restrictions are placed on $f'(z)$ in order to construct the C_n's.

To simplify analysis we will represent expression (2.16) in the general form

$$\frac{G_n(z)}{G_0(z)}=R_n(z)+C_nG_0(z)-H_n(z) \tag{2.25}$$

where all the functions on the right-hand side are analytic in the unit disc, and the integral function R_n has the property that $R_n'(\delta)=0$. From (2.15) it follows that

$$\frac{G_n'(z)}{G_0'(z)}=2R_n(z)+2C_nG_0(z)-H_n(z). \tag{2.26}$$

For $n=1$, substitute the expressions given by the right-hand side of (2.25) and (2.26) into the right-hand side of (2.13). Then, on combining terms, and using the result

$$\text{Real } H_1(z)=\frac{1}{|G_0(z)|^2}, \qquad |z|=1,$$

we obtain for $|z|=1$

$$h_2(z)=\text{Real }[C_1^2G_0^2+2C_1G_0R_1-2\bar{C}_1/G_0]+\tilde{h}_2(z)$$

where $\tilde{h}_2(z)$ represent terms which do not contain C_1.

Let $\tilde{H}_2(z)$ be a function which is analytic on the unit disc, such that on the boundary

$$\text{Real } \tilde{H}_2(z)=\tilde{h}_2(z), \qquad |z|=1.$$

In a manner similar to that which we used to obtain the coefficient of C_0 in $H_1(z)$, it can be shown that the analytic function

$$\frac{-2}{\sqrt{8}}[\bar{C}_1C_0+\bar{C}_1[P_0(z)-P_0(\delta)]/(z-\delta)+zC_1\overline{P_0}(\delta)/(1-z\bar{\delta})] \tag{2.27}$$

has the property that its real part is equal to Real $(-2\bar{C}_1/G_0)$ on the boundary $|z|=1$.

Since G_0, R_1 are analytic functions of z on the unit disc, it follows that $H_2(z)$ is a linear combination of $\tilde{H}_2(z)$, $C_1^2G_0^2+2C_1G_0R_1$ and expression (2.27). Hence we have

$$\sqrt{8}\,H_2'(\delta)\equiv-\bar{C}_1P_0''(\delta)+2\bar{f}'(\delta)C_1+\sqrt{8}\,\tilde{H}_2'(\delta)=0. \tag{2.28}$$

Thus as before, this equation can be solved for C_1 provided that $D(\delta)\neq 0$.

To find the remaining constants C_n $(n \geqq 2)$, we note that in a manner similar to that above, (2.14) can be reduced to the form

$$h_n(z) = \text{Real}\,[2C_{n-1}G_0(R_1 + C_1G_0) - 2\bar{C}_{n-1}/G_0] + \tilde{h}_n(z)$$

for $|z| = 1$, where $\tilde{h}_n(z)$ represent terms which do not contain the constant C_{n-1}. From this, one can show that (2.17) reduces to an equation of the form

$$-\bar{C}_{n-1}P_0''(\delta) + 2\bar{f}'(\delta)C_{n-1} + \sqrt{8}\,\tilde{H}_n'(\delta) = 0$$

which can be solved for C_{n-1}, provided that $D(\delta) \neq 0$.

We can now summarize the results as follows. Define the nth $(n \geqq 3)$ order solution by

$$\exp u_n = \frac{|1 - z\bar{\delta}|^4 |1 + 2\sum_{m=1}^{n-1} \lambda^m G_m/G_0 + \sum_{m=1}^{n-1} \lambda^m H_m|^2}{[(\lambda/8)|P(z)|^2 + |z-\delta|^2 |1 + \sum_{m=1}^{n-1} \lambda^m G_m/G_0|^2]^2}.$$

THEOREM. *If the transformation function $f'(z)$ satisfies the implicit conditions*

(i) $D(\delta) \neq 0$

(ii) *$P(z)$ has no zero in the closed unit disc,*

then the nth order solution u_n $(n \geqq 3)$ given by (2.29) *satisfies the equations*

$$-\Delta u_n = \lambda |f'(z)|^2 e^{u_n}, \qquad |z| < 1,$$
$$u_n = O(\lambda^n), \qquad |z| = 1.$$

The restriction that

$$P_0(z) + (z - \delta)C_0$$

where C_0 is given by (2.21), and $P_0(z)$ by (2.4) should have no zeros in the unit disc places an implicit condition on the class of transformations $f(z)$ and the root δ of (1.6) if there is more than one. It would be of future interest to be able to place the above restriction in a more explicit or transparent form. However for present purposes it is important to note that the set of transformations for which the above holds is not empty. This can be seen as follows. For the identity transformation where $f'(z) \equiv 1$, it can be shown that $P_0(z) \equiv -1$, and $C_0 = 0$. Hence it is obvious that perturbations of the identity,

$$f'(z) = 1 + \varepsilon g'(z)$$

where ε is sufficiently small, will satisfy the above restriction.

3.1. Newton process. In § 2, we constructed first and higher order asymptotic representations for the large norm (large maximum) solution to problem (P′). We want to examine whether these could be used to generate an exact solution using an iterative scheme like Newton's method. To simplify analysis we will consider here the modified's method.

Problem (P′) will be converted to the nonlinear integral

$$u(x_0) = \lambda \int_{|x| \leqq 1} g(x_0; x) |f'(z)|^2 e^{u(x)}\, dx, \tag{3.1}$$

where

$$g(x_0; x) = -\frac{1}{4\pi} \ln \left| \frac{z - z_0}{1 - \bar{z}_0 z} \right|^2, \tag{3.2}$$

with $z = x_1 + ix_2$, is the Green's function for the Dirichlet problem associated with the unit disc. For convenience of notation, rewrite the above equation in the following form

$$u = \mathbb{K}(u), \tag{3.3}$$

where $\mathbb{K}(u)$ is a nonlinear operator which maps the Banach space of continuous functions into itself. Hence in the following analysis the norm

$$\|u\| = \max_{|x| \leqq 1} |u(x)|$$

will be employed.

Let $\mathbb{K}'_u$ be the associated Fréchet derivative [7] of $\mathbb{K}(u)$, in which case it is the linear integral operator with kernel

$$\lambda g(x_0; x)|f'|^2 e^u.$$

In the iteration process below we will be using $\mathbb{K}'_{u_0}$ where u_0 is the initial approximation to be used in the iteration process.

Equation (3.3) is then re-written in the following form:

$$u = u - (\mathbb{I} - \mathbb{K}'_{u_0})^{-1}(u - \mathbb{K}(u)) = S(u)$$

where $S(u)$ represents a nonlinear operator which maps the Banach space of continuous functions into itself.

The modified Newton process is described by the iterative procedure

$$u_{n+1} = S(u_n), \qquad n = 0, 1, 2, \cdots, \tag{3.4}$$

starting from an initial approximation u_0. Conditions for the process to converge for a general class of nonlinear equations are given in Vainberg [7] or Rall [8].

Let

$$\|(\mathbb{I} - \mathbb{K}'_{u_0})^{-1}\mathbb{K}'_{u_0}\| \leqq \Gamma, \qquad (1+\Gamma)\|u_0 - \mathbb{K}(u_0)\| \leqq \varphi(0), \tag{3.5}$$

and

$$\varphi'(t) = \Gamma(e^t - 1).$$

It can then be shown that

$$\|S'(u)\| \leqq \varphi'(t), \quad \text{for } \|u - u_0\| \leqq t,$$

$$\|S(u_0) - u_0\| \leqq \varphi(0);$$

hence the operator $S(u)$ is majorized by $\varphi(t)$ where

$$\varphi(t) = \varphi(0) + \Gamma(e^t - 1 - t).$$

From Vainberg [7], the Newton process (3.4) will converge to u^* if the equation

$$\varphi(t) = t \tag{3.6}$$

has a positive solution, and $\|u^* - u_0\| \leqq t^*$ where t^* is the smallest positive root of (3.6). It can be shown that Equation (3.6) has a positive root t^* if

$$1 + \varphi(0) \leqq (1+\Gamma)\ln\left(\frac{1+\Gamma}{\Gamma}\right),$$

where

$$t^* \leqq \ln\left(\frac{1+\Gamma}{\Gamma}\right).$$

We can summarize the results as follows

THEOREM. *If*

$$\|u_0-\mathbb{K}(u_0)\|\leqq \ln\left(\frac{1+\Gamma}{\Gamma}\right)-(1+\Gamma)^{-1} \tag{3.7}$$

where Γ *is the upper bound in inequality* (3.5), *then the Newton process* (3.4) *will converge to a unique solution* u^* *such that*

$$\|u_0-u^*\|\leqq \ln\left(\frac{1+\Gamma}{\Gamma}\right).$$

In order to make use of the above theorem to find out whether the first or higher order approximations generated in § 2, can lead to the exact solution by use of the Newton process, we need to obtain estimates of the various quantities in inequality (3.7).

3.2. Estimate for $\|u_0-\mathbb{K}(u_0)\|$. As a first step we will obtain an estimate for $\|u_0-\mathbb{K}(u_0)\|$. From the differential equation (2.7) and the properties of the Green's function, it can be shown that

$$u_0-\mathbb{K}(u_0)=-\int_{|y|=1} u_0\frac{\partial g}{\partial n}\,ds. \tag{3.8}$$

Since $-\partial g/\partial n$ is positive and

$$\int_{|y|=1}\frac{\partial g}{\partial n}\,ds=-1,$$

it follows that

$$|u_0-\mathbb{K}(u_0)|\leqq \max_{|x|=1}|u_0(x)|. \tag{3.9}$$

Then if for u_0 we employ either the first, second or nth ($n\geqq 3$) order asymptotic approximation to u given by (2.6), (2.24), and (2.29) respectively, it follows from (3.9) and the properties of the solutions on the unit circle that

$$\|u_0-\mathbb{K}(u_0)\|\leqq \lambda^n M(\lambda), \tag{3.10}$$

where $M(\lambda)$ stands for a function bounded in λ for finite $\lambda\geqq 0$, and n refers to the order of the asymptotic expansion.

3.3. Estimate for Γ. An asymptotic estimate of $\|(I-\mathbb{K}'_{u_0})^{-1}\mathbb{K}'_{u_0}\|$ for $\lambda\to 0$ will be derived next. The operator $\mathbb{K}_{u_0}$ has the symmetrizable kernel

$$k(x_0,x)=g(x_0,x)\rho(x) \tag{3.11}$$

where $g(x_0,x)$, the Green's function associated with the Dirichlet problem for the unit disc centered at the origin, is positive and symmetric, and $\rho(x)$ is given by

$$\rho(x)=\lambda|f'(z)|^2e^{u_0(x)}. \tag{3.12}$$

Here $u_0(x)$ is the initial approximation in the Newton process, and is given by (2.6) or (2.24) or (2.29), as is indicated above.

As a first step we will obtain an estimate for $\|\mathbb{K}'_{u_0}\|$. Since the dominant contribution of the kernel comes from the neighborhood of the point $z=x_1+ix_2=\delta$, it will be convenient to decompose the unit disc into two parts D_1 and D_2 as follows. Define:

D_1, a disc with center δ and small but nonzero radius ε independent of λ.

D_2, the complement of D_1 with respect to $|x|<1$.

It follows from Appendix A that for $x \in D_1$,

$$\rho(x) \sim \rho_0(x)[1+O(\lambda^{1/2})],$$

where

$$\rho_0(x) = \frac{8\mu}{[\mu+|z-\delta|^2]^2}, \tag{3.13}$$

and

$$\mu = \frac{\lambda}{8}|P_0(\delta)|^2, \tag{3.14}$$

and for $x \in D_2$,

$$\rho(x) = \lambda |f'(z)|^2 \left|\frac{1-z\bar{\delta}}{z-\delta}\right|^4 [1+O(\lambda)].$$

From this, it follows that for λ sufficiently small

$$|k(x_0, x)| \leqq g(x_0, x)\rho_0(x)M_1, \qquad x \in D_1, \tag{3.15}$$

$$|k(x_0, x)| \leqq g(x_0, x)\lambda M_2, \qquad x \in D_2, \tag{3.16}$$

where M_1 and M_2 are appropriate constants. With these, we can show the following

LEMMA 1. *For λ sufficiently small,*

$$\|\mathbb{K}'_{u_0}\| \leqq \text{constant} \ln\left(\frac{1}{\lambda}\right).$$

Proof. From (3.15) and (3.16) it follows that

$$\int_{D_2} |k(x_0, x)|\, dx \leqq \text{constant}\ \lambda$$

$$\begin{aligned} \int_{D_1} |k(x_0, x)|\, dx &\leqq M_1 \int_{D_1} g(x_0, x)\rho_0(x)\, dx \\ &\leqq M_1 \int_{D_1} \left[-\frac{1}{2\pi}\ln|z-z_0| + M_3\right]\rho_0(x)\, dx, \end{aligned} \tag{3.17}$$

where the singular part is separated out of the expression for the Green's function given by (3.2) and M_3 is a bound on the remainder. The above expression can be reduced by using local polar coordinates (r, θ) centered at δ. Since $\rho_0(x)$ is independent of θ, the right-hand side of the above inequality reduces to

$$-M_1 \int_0^\varepsilon \ln(r_>)\rho_0 r\, dr + 2\pi M_1 M_3 \int_0^\varepsilon \rho_0 r\, dr$$

where $r_>$ is the max (r, r_0). It thus can be shown that

$$\int_{D_1} |k(x_0, x)|\, dx \leqq \begin{cases} M, & \text{for } x_0 \in D_2, \\ M - 2\ln(r_0^2+\mu), & \text{for } x_0 \in D_1, \end{cases} \tag{3.18}$$

where M is a suitable constant depending on ε. Since

$$\|\mathbb{K}'_{u_0}\| \leqq \max_{|x_0|<1} \int_{D_1+D_2} |k(x_0, x)|\, dx,$$

it follows that

$$\|\mathbb{K}'_{u_0}\| \leqq \text{constant} \ln\left(\frac{1}{\lambda}\right). \qquad \text{Q.E.D.}$$

We now want an estimate for the norm of the inverse of the operator $\mathbb{I}-\mathbb{K}'_{u_0}$. As will be seen this operator will become singular as $\lambda \to 0$; hence the analysis will not be straightforward. To obtain an estimate for the norm of the inverse operator, we will decompose the equation

$$(I-\mathbb{K}'_{u_0})v = w \tag{3.19}$$

into two equations, so that we can make use of the fact that the dominant contribution of the kernel comes from the neighborhood of the point $z=\delta$. To do so, define the following quantities

$$v_i(x)=\begin{cases} v(x), & x\in D_i,\\ 0, & \text{otherwise}\end{cases}$$

$$k_{ij}(x_0,x)=\begin{cases} k(x_0,x), & \text{for } x_0\in D_i,\ x\in D_j,\\ 0, & \text{otherwise}\end{cases}$$

where i, j take on the values 1 and 2. Thus we have

$$v=v_1+v_2, \qquad k(x_0,x)=\sum_{i,j=1}^{2} k_{ij}(x_0,x).$$

Let $\mathbb{K}_{ij}$ be the integral operator associated with kernel k_{ij}. Thus $\mathbb{K}_{ij}$ maps the Banach space of continuous functions defined on D_j, into the Banach space of continuous functions D_i. The norm of the operator $\mathbb{K}_{ij}$ will be defined appropriately

$$\|\mathbb{K}_{ij}\| = \sup_{v_j} \frac{\|\mathbb{K}_{ij}v_j\|}{\|v_j\|},$$

where

$$\|v_j\| = \max_{x\in D_j} |v_j(x)|, \qquad \|\mathbb{K}_{ij}v_j\| = \max_{x_0\in D_i} |\mathbb{K}_{ij}v_j|.$$

It follows from inequality (3.17), that for λ sufficiently small

$$\|\mathbb{K}_{22}\| \leqq M_{22}\lambda, \qquad \|\mathbb{K}_{12}\| \leqq M_{12}\lambda,$$

and from inequality (3.18)

$$\|\mathbb{K}_{21}\| \leqq M_{21},$$

where M_{ij} are appropriate constants.

Equation (3.19) can now be decomposed into the two equations

$$v_1-\mathbb{K}_{11}v_1-\mathbb{K}_{12}v_2=w_1, \tag{3.20a}$$

$$v_2-\mathbb{K}_{22}v_2-\mathbb{K}_{21}v_1=w_2. \tag{3.20b}$$

For λ sufficiently small $(I-\mathbb{K}_{22})^{-1}$ exists as a Neumann series; hence (3.20b) can be inverted to yield

$$v_2=(I-\mathbb{K}_{22})^{-1}w_2+(I-\mathbb{K}_{22})^{-1}\mathbb{K}_{21}v_1. \tag{3.21a}$$

This combined with (3.20a) yields

$$v_1-\tilde{\mathbb{K}}v_1=w_1+\mathbb{K}_{12}(I-\mathbb{K}_{22})^{-1}w_2 \tag{3.21b}$$

where the operator $\tilde{\mathbb{K}}$,

$$\tilde{\mathbb{K}}=\mathbb{K}_{11}+\mathbb{K}_{12}(I-\mathbb{K}_{22})^{-1}\mathbb{K}_{21}, \tag{3.22}$$

maps the Banach space of continuous functions defined on D_1, into itself.

We are now in a position to obtain the following intermediate result.

LEMMA 2. *For λ sufficiently small,*

$$\|(I-\mathbb{K}'_{u_0})^{-1}\| \leqq \max\,[m_1\|(I-\tilde{\mathbb{K}})^{-1}\|,\ m_2+m_3\|(I-\tilde{\mathbb{K}})^{-1}\|]$$

where m_1, m_2 and m_3 are appropriate constants.

Proof. From the decomposition of continuous functions $v(x)$ defined on $|x|<1$, into parts v_1 and v_2, it follows that

$$\|v\| = \max_{|x|<1} |v| = \max\,[\|v_1\|, \|v_2\|].$$

Then it follows that for the functions w_i used in (3.21a) and (3.21b)

$$\|w_i\| \leqq \|w\|.$$

Using this last result, and the fact that

$$\|(I-\mathbb{K}_{22})^{-1}\| \leqq (1-\lambda M_{22})^{-1}$$

it follows from (3.21b) that

$$\|v_1\| \leqq \|(I-\tilde{\mathbb{K}})^{-1}\|\{1+\lambda M_{12}(1-\lambda M_{22})^{-1}\}\|w\|,$$

and (3.21a) that

$$\|v_2\| \leqq (1-\lambda M_{22})^{-1}[\|w\|+M_{21}\|v_1\|].$$

Thus one obtains

$$\|v\| \leqq \max\,[m_1\|(I-\tilde{\mathbb{K}})^{-1}\|,\ m_2+m_3\|(I-\tilde{\mathbb{K}})^{-1}\|]\|w\|,$$

where the constants m_i are chosen so that for λ sufficiently small,

$$m_1 \geqq 1+\lambda M_{12}(1-\lambda M_{22})^{-1}, \qquad m_2 \geqq (1-\lambda M_{22})^{-1}, \qquad m_3 \geqq M_{21}(1-\lambda M_{22})^{-1} m_1.$$

From the relationship between v and w given by (3.19) the result for $\|(I-\mathbb{K}'_{u_0})^{-1}\|$ follows. Q.E.D.

The problem now reduces to finding an estimate for the norm of $(I-\tilde{\mathbb{K}})^{-1}$. To obtain this we will introduce a new integral operator $\mathbb{K}_0$ whose eigenvalues and eigenfunctions can be determined asymptotically and which, as will be shown, approximates the operator $\mathbb{K}_{11}$. The operator $\mathbb{K}_0$ will have kernel $k_0(x_0, x)$ given by

$$k_0(x_0, x) = G_0(x_0, x)\rho_0(x), \qquad x, x_0 \in D_1, \tag{3.23}$$

where

$$G_0 = -\frac{1}{4\pi} \ln \left|\frac{z-z_0}{1-|\delta|^2}\right|^2.$$

Since $\mathbb{K}_0$ is a symmetric nonnegative operator, its eigenvalues Λ_n are real and nonnegative. The corresponding eigenfunctions φ_n,

$$\Lambda_n \varphi_n = \mathbb{K}_0 \varphi_n,$$

form an orthonormal set in the Hilbert space with measure $\rho_0(x)\,dx$

$$\int_{D_1} \varphi_n \bar{\varphi}_m \rho_0(x)\,dx = \delta_{nm}.$$

The asymptotic behavior as $\lambda \to 0$ of Λ_n and φ_n are given in Appendix B. In particular it is shown that the largest eigenvalue Λ_1 is given by

$$\Lambda_1 \sim 4 \ln \left(\frac{1-|\delta|^2}{\sqrt{\mu}} \right)$$

and Λ_2, Λ_3 and Λ_4 approach unity as $\lambda \to 0$. The remaining eigenvalues tend to a number less than or equal to $1/3$ as $\lambda \to 0$.

Next we want to investigate the closeness of the approximation of $\mathbb{K}_0$ to $\mathbb{K}_{11}$ and hence $\tilde{\mathbb{K}}$. Recall that $\mathbb{K}_{11}$ is just the operator $\mathbb{K}'_{u_0}$ restricted to functions defined on D_1, hence has kernel $k(x_0, x)$, given by (3.11) where for $x \in D_1$,

$$\rho(x) \sim \rho_0(x)[1 + O(\lambda^{1/2})].$$

For further convenience we shall define the integral operator $\mathbb{K}_1$, with kernel $k_1(x_0, x)$ defined on $D_1 \times D_1$, given by

$$k_1(x_0, x) = \frac{1}{2\pi} \ln [|1 - \bar{z}_0 z|/(1-|\delta|^2)] \rho_0(x). \tag{3.24}$$

It then follows from (3.11), (3.23) and (3.24) together with the above comments, that

$$(\mathbb{K}_{11} - \mathbb{K}_0)v = \mathbb{K}_0 \left(\frac{\rho}{\rho_0} - 1 \right) v + \mathbb{K}_1 \frac{\rho}{\rho_0} v. \tag{3.25}$$

Using techniques similar to that in the proof of Lemma 1, one can show that for the operator with kernel $k_0(x_0, x)(\rho/\rho_0 - 1)$

$$\left\| \mathbb{K}_0 \left(\frac{\rho}{\rho_0} - 1 \right) \right\| \leqq O(\lambda^{1/2} \ln \lambda).$$

To obtain the estimate for the norm of $\mathbb{K}_1$, split its kernel into two parts

$$k_1(x_0, x) = \frac{1}{2\pi} \ln [|1 - \bar{z}_0 \delta|/(1-|\delta|^2)] \rho_0(x) + \frac{1}{2\pi} \ln \left| 1 - \frac{\bar{z}_0 r e^{i\theta}}{(1-\bar{z}_0 \delta)} \right| \rho_0(x). \tag{3.26}$$

It can be shown that on expanding the second term, and using the result that

$$\int_{D_1} r \rho_0(x)\, dx \leqq O(\lambda^{1/2}),$$

then

$$\|\mathbb{K}_1\| \leqq \max_{x_0 \in D_1} \frac{1}{2\pi} \int_{D_1} \left| \ln \left| \frac{1 - \bar{z}_0 \delta}{1 - |\delta|^2} \right| \right| \rho_0(x)\, dx + O(\lambda^{1/2})$$

$$\leqq 4 \ln \left[1 - \frac{\varepsilon |\delta|}{(1-|\delta|^2)} \right] + O(\lambda^{1/2}).$$

Thus for λ sufficiently small, we can choose ε small enough so that $\|\mathbb{K}_1\| \ll 1$, with the result that

$$\|\mathbb{K}_{11} - \mathbb{K}_0\| \ll 1.$$

However this result can be made stronger if one considers the second iterate. Since

$$\int_{D_1} \left| \ln \left| \frac{1 - \bar{z} \delta}{1 - |\delta|^2} \right| \rho_0(x) \right| dx \leqq O(\lambda^{1/2})$$

it can be shown that the norm of the second iterate of $\mathbb{K}_1$ is the order of $\lambda^{1/2}$, and it follows that

$$\|(\mathbb{K}_{11}-\mathbb{K}_0)^2\| \leqq O(\lambda^{1/2}\ln\lambda). \tag{3.27}$$

We are now in a position to consider $(I-\tilde{\mathbb{K}})^{-1}$. The operator given by (3.22) will be placed in the form

$$\tilde{\mathbb{K}} = \sum_{j=2}^{4} \Lambda_j \varphi_j(x_0)(\cdot, \varphi_j)_\rho + \mathbb{L} + \mathbb{N},$$

where

$$\mathbb{N} = \mathbb{K}_0 - \sum_{j=2}^{4} \varphi_j(x_0)(\cdot, \varphi_j)_\rho, \tag{3.28}$$

$$\mathbb{L} = \mathbb{K}_0\left(\frac{\rho}{\rho_0}-1\right) + \mathbb{K}_1\frac{\rho}{\rho_0} + \mathbb{K}_{12}(I-\mathbb{K}_{22})^{-1}\mathbb{K}_{21}, \tag{3.29}$$

and Λ_j, φ_j are the eigenvalues and eigenfunctions of $\mathbb{K}_0$, and

$$(u, v)_\rho = \int_{D_1} u\bar{v}\rho_0(x)\,dx.$$

Since the eigenvalues of $\mathbb{N}$ are bounded from unity as $\lambda \to 0$, the inverse of $(I-\mathbb{N})$ exists as a bounded operator for λ sufficiently small. We can now use Schmidt's method [9] for finding $(I-\tilde{\mathbb{K}})^{-1}$. The equation

$$(I-\tilde{\mathbb{K}})v = w$$

can be placed in the form

$$(I-(I-\mathbb{N})^{-1}\mathbb{L})v = \sum_{j=2}^{4} \Lambda_j (I-\mathbb{N})^{-1}\varphi_j(v, \varphi_j)_\rho + (I-\mathbb{N})^{-1}w. \tag{3.30}$$

Since $\mathbb{K}_{12}$ is the order of λ, the third term in expression (3.29) for $\mathbb{L}$ is small. From inequality (3.27) it then follows that

$$\|\mathbb{L}^2\| \leqq O(\lambda^{1/2}\ln\lambda).$$

It can be shown in a similar manner that

$$\|\mathbb{M}^2\| \leqq O(\lambda^{1/2}\ln\lambda),$$

where

$$\mathbb{M} = (I-\mathbb{N})^{-1}\mathbb{L}.$$

Thus the inverse of $(I-\mathbb{M})$ exists as a Neumann series, and hence equation (3.30) reduces to

$$v = \sum_{j=2}^{4} \Lambda_j (I-\mathbb{M})^{-1}\varphi_j(v, \varphi_j)_\rho + (I-\mathbb{M})^{-1}(I-\mathbb{N})^{-1}w, \tag{3.31a}$$

where use is made of the fact that $(I-\mathbb{N})^{-1}\varphi_j = \varphi_j$, for $j = 2, 3, 4$. The coefficients

$$c_j = \Lambda_j(v, \varphi_j)_\rho \tag{3.31b}$$

are found from the systems of equations

$$\sum_{j=2}^{4} a_{ij}c_j = -w_i, \qquad i = 2, 3, 4, \tag{3.32}$$

where

$$a_{ij}=((I-\mathbb{M})^{-1}\varphi_j,\varphi_i)_\rho-\delta_{ij}/\Lambda_i,\qquad w_j=((I-\mathbb{M})^{-1}(I-\mathbb{N})^{-1}w,\varphi_j)_\rho.$$

The problem of estimating $\|(I-\mathbb{K})^{-1}\|$ is reduced to finding the asymptotic solution of the linear system of equations (3.32). This requires evaluation of the terms

$$((I-\mathbb{M})^{-1}\varphi_j,\varphi_i)_\rho.$$

Since these terms involve an inner product defined on the Hilbert space of square integrable functions with measure $\rho_0(x)\,dx$ on the domain D_1, we can use the appropriate Hilbert space norm of the operators to obtain the required order estimate. In particular, it can be shown that the Hilbert space norm of $\mathbb{K}_1$, namely $\|\mathbb{K}_1\|_2$ is the order of $\lambda^{1/2}$, hence $\|\mathbb{M}\|_2$ is the order of $\lambda^{1/2}\ln\lambda$ at most. Thus we obtain the result

$$((I-\mathbb{M})^{-1}\varphi_j,\varphi_i)_\rho=(\varphi_j,\varphi_i)_\rho+(\mathbb{L}\varphi_j,\varphi_i)_\rho+(\mathbb{L}\mathbb{M}\varphi_j,\varphi_i)+O(\lambda^{1/2}\ln\lambda)^3$$

where use is made of the result

$$(\mathbb{M}\varphi_j,\varphi_i)_\rho=((I-\mathbb{N})^{-1}\mathbb{L}\varphi_j,\varphi_i)=(\mathbb{L}\varphi_j,\varphi_i)_\rho$$

where i belong to the set $\{2,3,4\}$.

In the following we shall neglect terms of order $\lambda\ln\lambda$ or higher when either $i=4$ or $j=4$ or both. When both i,j belong to the set $\{2,3\}$, we will require terms of order λ and neglect terms of order $\lambda/\ln\lambda$. With these requirements on the order of the desired approximation, it can be shown that when we substitute expression (3.29) for $\mathbb{L}$ into $(\mathbb{L}\varphi_j,\varphi_i)_\rho$ we can neglect the components arising from the third term on the right-hand side of (3.29). Hence for i,j belonging to the set $\{2,3,4\}$, we find

$$\begin{aligned}(\mathbb{L}\varphi_j,\varphi_i)_\rho&\sim\left(\mathbb{K}_1\frac{\rho}{\rho_0}\varphi_j,\varphi_i\right)_\rho+\left(\mathbb{K}_0\left(\frac{\rho}{\rho_0}-1\right)\varphi_j,\varphi_i\right)_\rho\\&\sim\left(\frac{\rho}{\rho_0}\varphi_j,\mathbb{K}_1\varphi_i\right)_\rho+\Lambda_i\left(\left(\frac{\rho}{\rho_0}-1\right)\varphi_j,\varphi_i\right)_\rho.\end{aligned}$$

For further reduction we require the following expansion for ρ/ρ_0 given in Appendix A,

$$\rho/\rho_0=1+\eta_1+\eta_2+\cdots,$$

where η_1, the term of order $\lambda^{1/2}$ is given explicitly by

$$\eta_1=-\frac{4\mu r}{(\mu+r^2)}q\cos(\theta+2q),$$

with $C_0/P_0(\delta)=q\,e^{i\alpha_q}$, and η_2 gives rise to terms of order λ or $\lambda\ln\lambda$.

In light of the above discussion on the required order of approximation, we find that when i,j belong to the set $\{2,3\}$, the required approximation is given by

$$(\mathbb{L}\varphi_j,\varphi_i)_\rho\sim(\varphi_j,\mathbb{K}_1\varphi_i)_\rho+\Lambda_i(\eta_1\varphi_j,\varphi_i)_\rho+(\eta_1\varphi_j,\mathbb{K}_1\varphi_i)_\rho+\Lambda_i(\eta_2\varphi_j,\varphi_i)_\rho.$$

However when either $i=4$ or $j=4$ or both, then we can neglect the last two terms on the right-hand side of the above expression.

When i,j belong to the set $\{2,3\}$ the evaluation of the additional term $(\mathbb{L}\mathbb{M}\varphi_j,\varphi_i)_\rho$ to order λ is required. For this, the operator $\mathbb{L}$ may be approximated by $\mathbb{K}_1+\mathbb{K}_0\eta_1$, in which case

$$\begin{aligned}(\mathbb{L}\mathbb{M}\varphi_j,\varphi_i)_\rho&\sim(\mathbb{M}\varphi_j,\mathbb{K}_1\varphi_i)_\rho+\Lambda_i(\eta_1\mathbb{M}\varphi_j,\varphi_i)\\&\sim(\mathbb{K}_1\varphi_j+\mathbb{K}_0\eta_1\varphi_j,(I-\mathbb{N})^{-1}\mathbb{K}_1\varphi_1)_\rho+\Lambda_i(\mathbb{K}_1\varphi_j+\mathbb{K}_0\eta_1\varphi_j,(I-\mathbb{N})^{-1}\eta_1\varphi_i)_\rho.\end{aligned}$$

It can be shown that this may be further simplified to

(3.33) $$(\mathbb{LM}\varphi_j, \varphi_i)_\rho \sim -(\eta_1\varphi_j, \varphi_1)_\rho(\varphi_1, \mathbb{K}_1\varphi_i + \eta_1\varphi_i)_\rho + O(\lambda \ln \lambda).$$

We are now in a position to evaluate a_{ij}, when either $i=4$ or $j=4$ or both. a_{ij} is given by

(3.34) $$a_{ij} = (1-1/\Lambda_i)\delta_{ij} + (\varphi_j, \mathbb{K}_1\varphi_i)_\rho + \Lambda_i(\varphi_j, \eta_1\varphi_i)_\rho.$$

Using the results of Appendix C, it can be shown that

$$a_{24} \sim a_{34} \sim a_{42} \sim a_{43} \sim O(\lambda^{1/2}/\ln \lambda)$$

and

$$a_{44} \sim (1-1/\Lambda_4) \sim \frac{3}{2}\frac{1}{\ln \mu},$$

where μ is given in terms of λ by (3.14). When i, j belong to the set $\{2,3\}$, expression (3.34) for a_{ij} must include the sum of the additional terms $(\eta_1\varphi_j, \mathbb{K}_1\varphi_i)_\rho$, $(\eta_2\varphi_j, \varphi_i)_\rho$ and the terms given by the right-hand side of (3.33). From this it can be shown (see remark at end of Appendix C)

$$a_{22} \sim a_{33} \sim -\mu\left[\frac{3}{(1-|\delta|^2)^2} + \frac{7}{15}q^2\right] + O(\lambda/\ln \lambda),$$

$$a_{23} \sim \overline{a_{32}} \sim -\mu\left[\frac{3}{2}p\, e^{i\alpha_p} + \frac{1}{15}q^2\, e^{i2\alpha_q}\right] + O(\lambda/\ln \lambda),$$

where α_p, α_q are defined by Equations (A.2) and (A.3).

Hence it can be seen, that except for the pathological case where

(3.35) $$a_{22}a_{33} - a_{23}a_{32} \sim O(\lambda/\ln \lambda)^2$$

the system of equations (3.32) yield solutions c_i, such that

$$|c_i| \leqq \text{constant } \lambda^{-1}\|w\|, \qquad i = 1, 2, 3.$$

We now have the intermediate result

LEMMA 3. *For λ sufficiently small*

$$\|(I-\tilde{\mathbb{K}})^{-1}\| \leqq \text{constant } (1/\lambda).$$

Proof. This follows directly from (3.31), the estimates on the $|c_i|$, and the fact that φ_i are bounded.

Finally we can now combine all the intermediate results, namely Lemma's 1, 2 and 3, to obtain

THEOREM. *For λ sufficiently small*

$$\Gamma = \|(I-\mathbb{K}'_{u_0})^{-1}\mathbb{K}'_{u_0}\| \leqq \text{constant } [1/\lambda \ln (1/\lambda)].$$

3.4. Convergence of the Newton process. We are now in a position to check inequality (3.7) to see if and when the Newton process generates the exact solution for λ sufficiently small.

For Γ large, inequality (3.7) can be written in the form

(3.36) $$\|u_0 - \mathbb{K}(u_0)\| \leqq \frac{1}{\Gamma(1+\Gamma)}.$$

From the above estimates, it is seen that the right-hand side of inequality (3.36) is the order of $(\lambda/\ln \lambda)^2$ for λ sufficiently small. Hence from the estimate given by inequality (3.10), it follows that inequality (3.7) will be satisfied when the nth order ($n \geqq 3$) approximation is used for u_0.

The results can be summarized as follows:

THEOREM. *If the transformation function $f'(z)$ and δ satisfy the implicit conditions*
(i) $D(\delta) \neq 0$
(ii) $P(z) = P_0(z) + C_0(z-\delta)$ *has no zeros on the unit disc where* C_0 *is given by* (2.21),
then in general[1] *for λ sufficiently small, the modified Newton iteration scheme converges to an exact solution of problem (P'), when the nth order $(n \geqq 3)$ asymptotic approximation given by* (2.29) *is used for the initial step.*

Conclusion. A prescription for the asymptotic expansion of the large norm solution (with single maximum) of Liouville's nonlinear differential equation has been given in terms of the parameter λ, which tends to zero. It is shown that under suitable conditions, the asymptotic expansion to at least the third order generates the exact solution through the application of the modified Newton process.

There still remains a considerable amount to be investigated, such as obtaining an explicit representation of the conditions (i) and (ii) of the Theorem in Section 3.4, in terms of the transformation function $f(z)$. It would also be useful to find out when large norm solutions of multiple maxima occur, since the deformation of the boundary from a circle could result in a shift or change from a large single maximum to a multiple maximum. Other questions to be considered are pointed out in § 1.

Appendix A. We want the asymptotic behavior of ρ/ρ_0 in the neighborhood of the point $z=\delta$, expressed in local polar coordinates (r, θ). We want only those terms that, when integrated from $r=0$ to $r=\varepsilon$, with measure $\rho_0(r) r\,dr$, yield expressions up to order λ. In the expression for $\rho(x)$ given by

$$\rho(x) = \lambda |f'(x)|^2 e^u$$

we will use the general nth order result for u given by (2.29).

Note that the expansion about the point $z=\delta$,

$$G_1/G_0 \sim G_1(\delta)/G_0(\delta) + O(r) \sim -H_1(\delta) + O(r),$$

yields

$$|1+2\lambda G_1/G_0 + \lambda H_1 + \cdots|^2 \sim |1+\lambda G_1/G_0 + \cdots|^2 \sim 1 - 2\lambda \text{ Real } H_1(\delta);$$

hence the expression for ρ reduces to

$$\rho \sim T_1/T_2, \tag{A.1}$$

where $T_1 = \lambda |f'(z)(1-z\bar{\delta})^2|^2 [1+2\lambda \text{ Real } H_1(\delta)]$, and $T_2 = [r^2 + \lambda/8 |P(z)|^2 [1+2\lambda \text{ Real } H_1(\delta)]]^2$. If we set

$$P_0''(\delta)/P_0(\delta) = p\, e^{i\alpha_p} \tag{A.2}$$

and use (1.6), it can be shown that

$$\lambda |f'(z)(1-z\bar{\delta})^2|^2 \sim 8\mu\{1 - r^2 p \cos(2\theta + \alpha_p)\},$$

where μ is given by (3.14).

[1] Note for the pathological case where the coefficients a_{ij} are such that Equation (3.35) holds, then an asymptotic approximation of order $n > 3$ will have to be used for the initial step.

If we set

(A.3) $$C_0/P_0(\delta)=q\,e^{i\alpha_q},$$

it can be shown that

$$\frac{\lambda}{8}|P(z)|^2\sim\mu\{1+2rq\cos(\theta+\alpha_q)+r^2p\cos(2\theta+\alpha_p)+r^2q^2+\cdots\}.$$

Insert this into the expression for T_2, then rewrite it in the form

$$T_2=(\mu+r^2)^2[1+T]^2$$

with the appropriate expression for T. With the expansion $(1+T)^{-2}=1-2T+3T^2+\cdots$, the expression for ρ can be reduced to

$$\rho/\rho_0=1+\eta_1+\eta_2+\cdots,$$

where

$$\eta_1=-\frac{4\mu rq}{(r^2+\mu)}\cos(\theta+\alpha_q),$$

$$\eta_2=-\left(\frac{\mu-r^2}{\mu+r^2}\right)2\lambda\ \mathrm{Real}\,H_1(\delta)+r^2q^2\left[\frac{6\mu^2}{(r^2+\mu)^2}-\frac{2\mu}{(r^2+\mu)}\right]$$
$$-p\cos(2\theta+\alpha_p)r^2\left[1+\frac{2\mu}{\mu+r^2}\right]+\frac{6\mu^2r^2q^2}{(\mu+r^2)^2}\cos(2\theta+2\alpha_q),$$

where ρ_0 is given by (3.13).

Appendix B. Here, the eigenvalues and eigenfunctions of the operator $\mathbb{K}_0$ are derived with emphasis placed upon their asymptotic behavior as $\lambda\to 0$.

With the use of local polar coordinates (r,θ) centered at δ, the region D_1 becomes the disc $0\leqq r\leqq\varepsilon$, $0\leqq\theta\leqq 2\pi$, and the kernel of operator $\mathbb{K}_0$, given by (3.23) becomes upon expanding the logarithmic term

$$k_0=\frac{8\mu}{(\mu+r^2)^2}\frac{1}{2\pi}\left\{\ln\left(\frac{1-|\delta|^2}{r_>}\right)+\sum_{n=1}^{\infty}\left(\frac{r_<}{r_>}\right)^n\frac{\cos n(\theta-\theta_0)}{n}\right\},$$

where $r_<=\min(r,r_0)$ and $r_>=\max(r,r_0)$. Since it is obvious that the eigenfunctions will have the separable form

$$\Phi=e^{\pm im\theta}\Psi_m(r),$$

the eigenvalue equation $\Lambda\Phi=\mathbb{K}_0\Phi$ reduces to the following

$$\Lambda\Psi_m(r_0)=\frac{1}{m}\int_0^{\varepsilon}\frac{4\mu}{(\mu+r^2)^2}\left(\frac{r_<}{r_>}\right)^m\Psi_m(r)r\,dr,\qquad m=1,2,3,\cdots,$$

$$\Lambda\Psi_0(r_0)=\int_0^{\varepsilon}\frac{8\mu}{(\mu+r^2)^2}\ln\left(\frac{1-|\delta|^2}{r_>}\right)\Psi_0(r)r\,dr,\qquad m=0.$$

The explicit results for the radial part of the eigenfunctions are obtained by differentiating the above equation twice to obtain a second order ordinary differential equation. Before proceeding in this manner it is best to first change the variable r to ξ where

(B.1) $$\xi=\frac{\mu-r^2}{\mu+r^2}.$$

Since

$$d\xi = -\frac{4\mu r\, dr}{(\mu + r^2)^2},$$

the above integral equation for Ψ_m becomes

$$\text{(B.2)} \qquad \Lambda\Psi_m(\xi_0) = \frac{1}{m}\int_{\xi_\varepsilon}^{1}\left(\frac{1-\xi_>}{1+\xi_>}\right)^{m/2}\left(\frac{1+\xi_<}{1-\xi_<}\right)^{m/2}\Psi_m(\xi)\, d\xi,$$

$$\text{(B.3)} \qquad \Lambda\Psi_0(\xi_0) = \int_{\xi_\varepsilon}^{1}\ln\left[\frac{(1-|\delta|^2)^2(1+\xi_<)}{\mu(1-\xi_<)}\right]\Psi_0(\xi)\, d\xi,$$

where ξ_ε is the value of ξ corresponding to $r = \varepsilon$.

The corresponding second order ordinary differential equation for Ψ_m in the variable ξ is now obtained in the manner indicated above. The resulting equation is recognizable as the associated Legendre equation. With the boundary condition that at $\xi = 1$ $(r = 0)$ Ψ_m must be finite, it follows that

$$\text{(B.4)} \qquad \Psi_m(\xi) = \mathcal{P}_\nu^m(\xi), \qquad \Lambda = \frac{2}{\nu(\nu+1)},$$

where $\mathcal{P}_\nu^m(\xi)$ is the associated Legendre function of the first kind.

To determine the value of ν, the boundary condition at ξ_ε is required. This is obtained by substituting back into (B.2) and (B.3) expression (B.4) for Ψ_m, and then using the following integrals:

$$\int^{\xi}\left(\frac{1-\xi}{1+\xi}\right)^{\pm m/2}\mathcal{P}_\nu^m(\xi)\, d\xi = \frac{1}{\nu(\nu+1)}\left(\frac{1-\xi}{1+\xi}\right)^{\pm m/2}\left[\mp m\mathcal{P}_\nu^m(\xi) - (1-\xi^2)\frac{d\mathcal{P}_\nu^m}{d\xi}\right],$$

$$\int^{\xi}\ln\left(\frac{1+\xi}{1-\xi}\right)\mathcal{P}_\nu(\xi)\, d\xi = \frac{1}{\nu(\nu+1)}\left[2\mathcal{P}_\nu(\xi) - \ln\left(\frac{1+\xi}{1-\xi}\right)(1-\xi^2)\frac{d\mathcal{P}_\nu}{d\xi}\right].$$

On equating both sides of the resulting equation, one ends with the results

$$\text{(B.5)} \qquad m\mathcal{P}_\nu^m(\xi_\varepsilon) - (1-\xi_\varepsilon^2)\frac{d\mathcal{P}_\nu^m}{d\xi_\varepsilon} = 0,$$

$$\text{(B.6)} \qquad \mathcal{P}_\nu(\xi_\varepsilon) - \ln\left(\frac{1-|\delta|^2}{\varepsilon}\right)(1-\xi_\varepsilon^2)\frac{d\mathcal{P}_\nu}{d\xi_\varepsilon} = 0.$$

These expressions are the required boundary conditions at $\xi = \xi_\varepsilon$ and hence will yield the appropriate values of ν.

We want the asymptotic values of ν and Λ as $\lambda \to 0$. Since

$$\xi_\varepsilon = \frac{\mu - \varepsilon^2}{\mu + \varepsilon^2} \sim -1 + \frac{2\mu}{\varepsilon^2}\cdots,$$

we see that $\xi_\varepsilon \to -1$ as $\lambda \to 0$. Thus we need to use the following asymptotic form for the Legendre function as $\xi \to -1$,

$$\mathcal{P}_\nu^m(\xi) \sim \frac{\cos(\pi\nu)}{m!}\left(\frac{1+\xi}{2}\right)^{m/2}\frac{\Gamma(\nu+m+1)}{\Gamma(\nu-m+1)}\left[1 + \left(\frac{m}{2} - \frac{\nu^2+\nu}{m+1}\right)\left(\frac{1+\xi}{2}\right)\right] - \frac{\sin(\pi\nu)\Gamma(m)2^{m/2}}{\pi(1+\xi)^{m/2}},$$

$$\mathcal{P}_\nu(\xi) \sim \cos(\pi\nu) - \frac{\sin(\pi\nu)}{\pi}\left[\ln\left(\frac{2}{1+\xi}\right) - 2\gamma - 2\psi(\nu+1)\right],$$

where ψ is the logarithmic derivative of the gamma function and γ is Euler's constant.

These results for $\xi \to -1$ follows directly from the relation

$$\mathcal{P}_\nu^m(\xi) = (-1)^m \mathcal{P}_\nu^m(-\xi)\cos(\pi\nu) - \frac{2}{\pi}(-1)^m \mathcal{Q}_\nu^m(-\xi)\sin(\pi\nu)$$

between the associated Legendre functions of the first and second kinds, and their asymptotic behavior as their arguments approach unity, given in [10].

It can thus be shown that boundary conditions yield the following asymptotic values of ν:

(B.7) $$\nu = n + \left(\frac{\mu}{\varepsilon^2}\right)^{m+1} \frac{(n+m)!}{(n-m)!} \frac{n(n+1)}{(m+1)!m!}, \qquad m = 1, 2, 3, \cdots,$$

(B.8) $$\nu = n + \left[\ln\left(\frac{1-|\delta|^2}{\sqrt{\mu}}\right) - 2\gamma - 2\psi(n+1)\right]^{-1}, \qquad m = 0,$$

where n takes on the values, m, $m+1$, $m+2$, $\cdots$.

We want to normalize the eigenfunction so that

$$(\Phi, \Phi)_\rho = \int_0^\varepsilon \int_0^{2\pi} |\Phi|^2 \rho_0(r) r\,dr\,d\theta = 1.$$

Set for $m = 1, 2, 3, \cdots$

(B.9) $$\Phi_n^{m\pm} = \alpha_\nu^m \mathcal{P}_\nu^m(\xi)\, e^{\pm im\theta},$$

where $\nu = \nu(n, m)$ is given by (B.7), and for $m = 0$,

(B.10) $$\Phi_n = \alpha_\nu \mathcal{P}_\nu(\xi),$$

where $\nu = \nu(n)$ is given by (B.8).

Thus we obtain for $m = 0, 1, 2, \cdots,$

(B.11) $$\alpha_\nu^m = \left[4\pi \int_{\xi_\varepsilon}^1 (\mathcal{P}_\nu^m(\xi))^2\,d\xi\right]^{-1/2},$$

which has the following asymptotic behavior as $\lambda \to 0$:

$$(\alpha_\nu^m)^2 \sim \frac{(n+1/2)(n-m)!}{4\pi(n+m)!}.$$

For the main body of the paper we want to order the eigenvalues in decreasing size, $\Lambda^1 \geqq \Lambda^2 \geqq \Lambda^3 \geqq \cdots$. With this convention we shall let the corresponding eigenfunction be φ', φ^2, etc. It follows that the four largest eigenvalues (as $\lambda \to 0$) and corresponding eigenfunctions are given by

$$\Lambda^1 \sim 4\ln\left(\frac{1-|\delta|^2}{\sqrt{\mu}}\right), \qquad \varphi' = \Phi_0,$$

$$\Lambda^2 \sim 1 - 3\left(\frac{\mu}{\varepsilon^2}\right)^2, \qquad \varphi^2 = \Phi_1^{1+},$$

$$\Lambda^3 = \Lambda^2, \qquad \varphi^3 = \Phi_1^{1-},$$

$$\Lambda^4 \sim 1 - \frac{3}{4}\left(\ln\frac{1-|\delta|^2}{\sqrt{\mu}}\right)^{-1}, \qquad \varphi^4 = \Phi_1.$$

Appendix C. We want to evaluate the coefficients $(\varphi_j, \mathbb{K}_1\varphi_i)_\rho$ and $(\varphi_j, \eta_1\varphi)_\rho$ where i, j are in the set $\{1, 2, 3, 4\}$. First note that the only nonzero terms of the first set of coefficients are given by

$$(\varphi_k, \mathbb{K}_1\varphi_2)_\rho = \overline{(\varphi_k, \mathbb{K}_1\varphi_3)}_\rho = \overline{(\varphi_2, \mathbb{K}_1\varphi_k)}_\rho = (\varphi_3, \mathbb{K}_1\varphi_k)_\rho,$$

where $k = 1$ or 4, and

$$(\varphi_2, \mathbb{K}_1\varphi_2)_\rho = \overline{(\varphi_3, \mathbb{K}_1\varphi_3)}_\rho.$$

Thus we need to compute only the three terms $(\varphi_k, \mathbb{K}_1\varphi_2)_\rho$ for $k = 1$, 2, and 4. Since the kernel of operator $\mathbb{K}_1$ can be expanded in the form

$$k_1(x_0, x) = \left\{\ln\left|\frac{1-\bar{z}_0 z}{1-|\delta|^2}\right| - \sum_{m=1} \frac{1}{m}\operatorname{Real}\left(\frac{\bar{z}_0 r\, e^{i\theta}}{1-\bar{z}_0\delta}\right)^m\right\}\frac{\rho_0(x)}{2\pi},$$

it follows that

$$\mathbb{K}_1\varphi_2 = -\mu^{1/2}\alpha_\nu^1\left(\frac{z_0}{1-z_0\bar{\delta}}\right)\int_{\xi_\varepsilon}^1 \left(\frac{1-\xi}{1+\xi}\right)^{1/2}\mathcal{P}_\nu^1(\xi)\,d\xi,$$

where α_ν^1 is the normalization constant associated with φ_2, and ν is given by (B.7) with $n = 1$. From the integral relation in the Appendix B and the eigenvalue equation (B.5) it follows

$$\mathbb{K}_1\varphi_2 = -\mu^{1/2}\left(\frac{z_0}{1-z_0\bar{\delta}}\right)\alpha_\nu^1\Lambda_2\left(\frac{1-\xi_\varepsilon}{1-\xi_\varepsilon}\right)^{1/2}\mathcal{P}_\nu^1(\xi_\varepsilon) \sim \frac{\mu^{1/2}}{2}\left(\frac{z_0}{1-z_0\bar{\delta}}\right)\sqrt{\frac{3}{\pi}}.$$

With the expansion

$$z_0(1-z_0\bar{\delta})^{-1} = \delta(1-|\delta|^2)^{-1} + r\,e^{i\theta}(1-|\delta|^2)^{-2} + \cdots$$

it follows immediately that $(\varphi_3, \mathbb{K}_1\varphi_2)_\rho = 0$, and

$$(\varphi_2, \mathbb{K}_1\varphi_2)_\rho \sim \frac{3}{2}\frac{\mu}{(1-|\delta|^2)^2}\int_{\xi_\varepsilon}^1\left(\frac{1-\xi}{1+\xi}\right)^{1/2}\mathcal{P}_\nu^1(\xi)\,d\xi$$

$$\sim \frac{-3\mu}{(1-|\delta|^2)^2}.$$

In a similar manner it can be shown that

$$(\varphi_1, \mathbb{K}_1\varphi_2)_\rho \sim \frac{\bar{\delta}}{(1-|\delta|^2)}\sqrt{6\mu},$$

$$(\varphi_4, \mathbb{K}_1\varphi_2)_\rho \sim \frac{\bar{\delta}}{(1-|\delta|^2)}\frac{3\sqrt{\mu/2}}{\ln\mu}.$$

The only nonzero components of $(\varphi_j, \eta_1\varphi_i)_\rho$ are

$$(\varphi_k, \eta_1\varphi_2)_\rho = \overline{(\varphi_k, \eta_1\varphi_3)}_\rho = \overline{(\varphi_2, \eta_1\varphi_k)}_\rho = (\varphi_2, \eta_1\varphi_k)_\rho,$$

where k takes on the values 1 and 4. These are given explicitly by

$$(\varphi_k, \eta_1\varphi_2)_\rho = -4\sqrt{\mu}\,\alpha_\nu^0\alpha_\nu^1 q\,e^{i\alpha_q}\pi\int_{\xi_\varepsilon}^1 \mathcal{P}_\nu(\xi)\mathcal{P}_{\nu'}^1(\xi)\sqrt{1-\xi^2}\,d\xi,$$

$$(\varphi_1, \eta_1\varphi_2)_\rho \sim \sqrt{\frac{2\mu}{3}}\,q\,e^{i\alpha_q}, \qquad (\varphi_4, \eta_1\varphi_2)_\rho \sim O(\lambda/\ln\lambda).$$

Remark. In the calculation of coefficients of the type $(\varphi_j, \eta_2\varphi_i)_\rho$ where i, j belong to the set $\{2, 3\}$, one can use the following asymptotic approximation of φ_2 and φ_3:

$$\bar{\varphi}_3 = \varphi_2 = -\frac{1}{4}\sqrt{\frac{3}{\pi}}\sqrt{1-\xi^2}\, e^{i\theta}.$$

Acknowledgment. The author wishes to thank the editor for helpful suggestions in re-writing the manuscript.

REFERENCES

[1] C. BANDLE, *Existence theorems, Qualitative results and a priori bounds for a class of non-linear Dirichlet problems*, Arch. Rational Mech. Anal., 58 (1975), pp. 219–238.

[2] M. G. CRANDALL AND P. N. RABINOWITZ, *Some continuation and variational methods for positive solutions of non-linear elliptic eigenvalue problems*, Ibid., 58 (1975), pp. 207–218.

[3] H. B. KELLER AND D. S. COHEN, *Some positone problems suggested by non-linear heat generation*, J. Math. Mech., 16 (1967), pp. 1361–1376.

[4] T. LAETSCH, *On the number of solutions of boundary value problems with convex non-linearities*, J. Math. Anal. Appl., 35 (1971), pp. 389–404.

[5] J. LIOUVILLE, *Sur l'équation aux différences partielles* $\frac{\partial^2 \log \lambda}{\partial u\, \partial v} \pm \lambda/2a^2 = 0$, J. de. Math., 18 (1853), pp. 71–72.

[6] A. R. FORSYTH, *Partial Differential Equations*, vols. 5 and 6, Dover, New York, 1959.

[7] M. M. VAINBERG, *Variational Methods for the Study of Nonlinear Operators*, Holden-Day, San Francisco, 1964.

[8] L. B. RALL, *Nonlinear Functional Analysis and Applications*, Academic Press, New York, 1971.

[9] R. COURANT AND D. HILBERT, *Methods of Mathematical Physics*, vol. 2, Wiley Interscience, New York, 1966.

[10] A. ERDÉLYI, W. MAGNUS, F. OBERHETTINGER, AND F. TRICOMI, *Higher Transcendental Functions*, vol. 1, McGraw-Hill, New York, 1953.

SIAM J. MATH. ANAL.
Vol. 9, No. 6, December 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0005 $01.00/0

A VARIATIONAL APPROACH TO MULTI-PARAMETER EIGENVALUE PROBLEMS IN HILBERT SPACE*

PAUL BINDING† AND PATRICK J. BROWNE‡

Abstract. Let T_r and V_{rs} be self-adjoint linear operators on Hilbert spaces H_r, $1 \leqq r \leqq k$. Assuming that T_r have compact resolvents and V_{rs} are bounded, we use standard variational characterizations of eigenvalues to treat the multiparameter case $(T_r + \sum_{s=1}^k \lambda_s V_{rs})x_r = 0$, $1 \leqq r \leqq k$. In particular we establish existence of a "purely point spectrum" satisfying cone monotonicity conditions in $\mathbb{R}^k$. We prove continuous, monotonic and Lipschitz parametric dependence theorems, and we examine $\mathbb{R}^k$-valued generalized Rayleigh quotients. In particular, we interpret the (not necessarily closed) convex hull of the spectrum as the "vectorial range" of these quotients.

1. Introduction. Atkinson's paper [3] which appeared in 1968 provided a stimulus for renewed interest in multiparameter spectral theory. This theory is now being investigated in abstract settings by several authors and has as its prime motivating examples eigenfunction expansions associated with linked systems of second order ordinary differential equations. These systems arise from the separation of variables for partial differential equations. Such problems have a long history and arise for example in considerations of the oscillations of circular and elliptic membranes. Mathieu functions are naturally connected with these problems. We mention but one reference, viz. [1]. Arscott's work [2] involving a single second order differential equation with several parameters is another interesting aspect of this theory. The survey book [1] mentioned above outlines further applications of multiparameter theory and has an extensive bibliography.

Our earlier work [7] discussed multiparameter problems for matrices, but with the differential equation applications in mind, it becomes necessary to investigate the theory in an infinite dimensional setting. The results presented here are so framed as to be directly applicable to multiparameter problems involving linked systems of ordinary differential equations over finite intervals.

We propose investigating certain classes of multi-parameter eigenvalue problems involving self-adjoint operators in general complex Hilbert spaces. This will extend our earlier finite dimensional results and it will be helpful to be familiar with at least the introduction of [7]. The conditions placed on our operators are of three types, two of which are most easily seen when $k = 1$; the problem then being to solve

$$(T + \lambda V)x = 0$$

for $\lambda \in \mathbb{R}$ and x in our Hilbert space H.

Our first assumption on this problem is that V is Hermitian and strongly negative definite in the sense that there is a constant $\alpha > 0$ such that

$$(Vy, y) \leqq -\alpha \|y\|^2, \quad \forall y \in H.$$

* Received by the editors November 11, 1976, and in final revised form February 22, 1978.

† Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta, Canada T2N 1N4. The work of this author was supported in part by the National Research Council of Canada under Grant A9071.

‡ Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta, Canada T2N 1N4. The work of this author was supported in part by the National Research Council of Canada under Grant A9073.

Secondly, we assume that T is self-adjoint with compact resolvent and is bounded below in the sense that there is a constant M such that

$$(Tx, x) \geqq M\|x\|^2, \quad \forall x \in \mathscr{D}(T).$$

The upshot of these conditions, as we shall show, is that for each $\mu \in \mathbb{R}$, $T+\mu V$ has a spectrum consisting entirely of eigenvalues of which at most a finite number are negative. Further, the lower eigenvalues of $T+\mu V$ are positive (negative) as $\mu \to -\infty(+\infty)$. These somewhat weaker statements could have been used in lieu of the assumptions listed above.

These hypotheses apply to a wide range of differential eigenvalue problems as well as to the matrix case treated in [7]. In particular, they lead to the existence of an increasing sequence λ^i of eigenvalues accumulating at no finite point. Also λ^i depends continuously on a parameter ν if V does as well. Such statements are standard for the one parameter problem and are contained in [7] for the matrix multi-parameter case.

The third assumption involves the multi-parameter structure and may be compared with the definiteness condition used in [7]. We delay a discussion of the condition until the problem is posed more precisely. However, it suffices to say that we shall investigate the lattice structure of the eigenvalues $\boldsymbol{\lambda} \in \mathbb{R}^k$ under the coordinate-wise partial order.

The paper is organized as follows. The abstract problem is posed in § 2 and the existence of eigenvalues established in § 3. In § 4 the lattice structure of the spectrum is investigated while in § 5 we extend the results of [7] concerning the vectorial range. In § 6 we investigate the situation in which the operators involved are allowed to depend on a parameter and we close with some examples.

2. The abstract problem and preliminary transformations. We are given Hilbert spaces H_r with unit spheres

$$S_r = \{u_r \in H_r \mid \|u_r\| = 1\}, \qquad 1 \leqq r \leqq k.$$

Self-adjoint operators in each space H_r are also given as follows:

(i) $V_{rs}: H_r \to H_r$ is bounded, $1 \leqq s \leqq k$,

(ii) $T_r: \mathscr{D}(T_r) \subset H_r \to H_r$ has compact resolvent and is bounded below; i.e. there is a constant α_r such that

$$(T_r u_r, u_r) \geqq \alpha_r \|u_r\|^2, \quad \forall u_r \in \mathscr{D}(T_r).$$

We shall find it convenient, for $u_r \in S_r$, $1 \leqq r \leqq k$, to use the notation $v_{rs}(u) = (V_{rs}u_r, u_r)$ to define $n \times n$ matrices

$$V^n(u) = [v_{rs}(u)]_{r,s=1}^n, \qquad 1 \leqq n \leqq k. \tag{1}$$

It is customary in multi-parameter theory to make a definiteness assumption on the V_{rs}: in our notation we assume

$$\gamma = \inf\{\det V^k(u) \mid u_r \in S_r\} > 0. \tag{2}$$

Our object is to find eigenvalues $\boldsymbol{\lambda} \in \mathbb{R}^k$ and eigenvectors $x_r \in S_r \cap \mathscr{D}(T_r)$ for the multi-parameter problem

$$W_r(\boldsymbol{\lambda})x_r = 0 \quad \text{where } W_r(\boldsymbol{\lambda}) = T_r + \sum_{s=1}^k \lambda_s V_{rs}, \quad 1 \leqq r \leqq k.$$

The first step in the analysis is to produce a linear transformation A on $\mathbb{R}^k$ so that $\boldsymbol{\lambda} \to \tilde{\boldsymbol{\lambda}} = A^{-1}\boldsymbol{\lambda}$ and $V_{rs} \to \tilde{V}_{rs} = \sum_{j=1}^k V_{rj}a_{js}$ and so that the resulting matrices $\tilde{V}^n(u)$ have determinants of fixed sign.

THEOREM 1. *There is a nonsingular linear transformation* A: $\mathbb{R}^k \to \mathbb{R}^k$ *so that if* $\tilde{\boldsymbol{\lambda}} = A^{-1}\boldsymbol{\lambda}$ *and* $\tilde{V}_{rs} = \sum_{j=1}^{k} V_{rj}a_{js}$ *then* $W_r(\boldsymbol{\lambda}) = T_r + \sum_{s=1}^{k} \tilde{\lambda}_s \tilde{V}_{rs}$ *and* $\inf\{(-1)^n \det \tilde{V}^n(u) \mid u_r \in S_r\} > 0$.

Proof. Fix $y_r \in S_r$, $1 \leqq r \leqq k$, and define the (r, s) element of A^{-1} by

$$[A^{-1}]_{rs} = -v_{rs}(y)$$

using the notation of (1), so A^{-1} is invertible by (2) and also

$$\tilde{v}_{rs}(y) = -\delta_{rs}.$$

Now choose u_r arbitrarily from S_r if $1 \leqq r \leqq n$ and $u_r = y_r$ if $n < r \leqq k$ to give

$$\begin{aligned}(-1)^n \det \tilde{V}^n(u) &= (-1)^k \det \tilde{V}^k(u)\\ &= (-1)^k \det V^k(u) \det A\\ &= \det V^k(u)/\det V^k(y).\end{aligned}$$

The result now follows directly from (2).

We conclude this section with some terminology and notation to be used at various subsequent points. If A is a linear operator on a Hilbert space H with identity map I then $A > 0$ means that A is positive definite on H. A is *strongly positive definite* ($A \gg 0$) if $A - \beta I > 0$ for some real $\beta > 0$ and if $A = A(\lambda)$ depends on a real parameter λ then A is *strongly increasing* in λ if for some real $\beta > 0$

$$\frac{A(\lambda) - A(\mu)}{\lambda - \mu} - \beta I > 0$$

whenever $\lambda \neq \mu$.

3. Existence of the spectrum. We assume for this section that the transformation of Theorem 1 has been carried out and for convenience we shall drop the tildas. The first task is to show that our operators $W_r(\boldsymbol{\lambda})$ obey the minimax principle for eigenvalues.

LEMMA 1. *Let A and B be linear operators on a Hilbert space H with* A^{-1} *compact and B bounded. Then if* $(A + B)^{-1}$ *is bounded, it is compact.*

Proof. Let $y_n \in H$ converge weakly to zero, so if $z_n = (A + B)^{-1}y_n$ then z_n also converges weakly to zero. Further we have $z_n = A^{-1}y_n - A^{-1}Bz_n$. This shows that z_n converges strongly to zero and so completes the proof.

COROLLARY. *For each* $\boldsymbol{\lambda} \in \mathbb{R}^k$, $W_r(\boldsymbol{\lambda})$ *has compact resolvent.*

Proof. Note that by hypothesis T_r has compact resolvent. Now apply the lemma.

We see now that for $\boldsymbol{\lambda} \in \mathbb{R}^k$, each $W_r(\boldsymbol{\lambda})$ is a self-adjoint operator bounded below and with compact resolvent. As such, $W_r(\boldsymbol{\lambda})$ has a spectrum consisting entirely of eigenvalues

$$\rho_r^0(\boldsymbol{\lambda}) \leqq \rho_r^1(\boldsymbol{\lambda}) \leqq \rho_r^2(\boldsymbol{\lambda}) \leqq \cdots$$

each of finite multiplicity and accumulating at no finite point. The minimax principle [8, p. 24] gives

(3) $\rho_r^i(\boldsymbol{\lambda}) = \max\{\min\{(W_r(\boldsymbol{\lambda})u_r, u_r) \mid u_r \in S_r \cap \mathscr{D}(T_r), (u_r, y_j) = 0\} \mid y_j \in H_r, 1 \leqq j \leqq i\}.$

We now present our principal existence theorem.

THEOREM 2. *Corresponding to each multi-index* $\mathbf{i} = (i_1, \cdots, i_k)$ *where each* $i_r \geqq 0$ *is an integer, there is an eigenvalue* $\boldsymbol{\lambda}^{\mathbf{i}}$ *and eigenvector* $x_r^{\mathbf{i}}$, $1 \leqq r \leqq k$, *so that* $\rho_r^{i_r}(\boldsymbol{\lambda}^{\mathbf{i}}) = 0$ *and* $W_r(\boldsymbol{\lambda}^{\mathbf{i}})x_r^{\mathbf{i}} = 0$, $1 \leqq r \leqq k$.

Proof. We shall argue by induction and prove slightly more than is stated. We assume that each T_r is replaced by $T_r + Q_r(\nu)$ where Q_r is bounded and depends continuously (in norm) on a Euclidean parameter ν and shall prove continuous dependence of $\boldsymbol{\lambda}^{\mathbf{i}}$ on ν as well.

Starting with $k = 1$, and suppressing subscripts, we are given

$$W(\lambda, \nu) = T + \lambda V + Q(\nu)$$

and $V \ll 0$ for each ν so $W(\lambda, \nu)$, and hence $\rho^i(\lambda, \nu)$ (by the minimax principle), strongly decrease in λ for each ν. Thus there is $\lambda^i(\nu)$ for which $\rho^i(\lambda^i(\nu), \nu) = b0$ and it remains to establish continuity of $\lambda^i(\nu)$ in ν.

Let $\nu_j \to \nu_*$ and restrict attention to a compact set of ν values G containing the ν_j and ν_*. By continuity of Q in ν, there is $\beta > 0$ *independent* of $\nu \in G$ so that

$$|\rho^i(\lambda, \nu) - \rho^i(\mu, \nu)| > \beta|\lambda - \mu|. \tag{4}$$

Setting $\lambda^i(\nu_j) = \lambda^i_j$ etc., we obtain

$$\begin{aligned}\beta|\lambda^i_j - \lambda^i_*| &< |\rho^i(\lambda^i_j, \nu_j) - \rho^i(\lambda^i_*, \nu_j)| \\ &= |\rho^i(\lambda^i_*, \nu_*) - \rho^i(\lambda^i_*, \nu_j)| \to 0 \quad \text{as } j \to \infty\end{aligned}$$

since ρ^i is continuous in ν by the minimax principle. It follows that $\lambda^i_j \to \lambda^i_*$ as required.

We turn now to the inductive step from $k-1$ to k, so for each λ_k we assume the existence of $\lambda^{\mathbf{i}}_r$, $1 \leqq r < k$, continuously dependent on ν *and* λ_k. (Technically we have enlarged the parameter space but the proof above is obviously dimension-free.)

For two different values λ_k and λ'_k we write $\lambda^{\mathbf{i}}_r(\lambda_k) = \lambda_r$, $\lambda^{\mathbf{i}}_r(\lambda'_k) = \lambda'_r$ and set $\mu_r = \lambda_r - \lambda'_r$. For $1 \leqq r < k$ let y_r be the minimizer of $(W_r(\boldsymbol{\lambda})u_r, u_r)$ over $u_r \in S_r$ and orthogonal to the first i_r eigenvectors of $W_r(\boldsymbol{\lambda}')$. From the minimax principle we obtain

$$(W_r(\boldsymbol{\lambda})y_r, y_r) \leqq 0 \leqq (W_r(\boldsymbol{\lambda}')y_r, y_r), \qquad 1 \leqq r < k.$$

Likewise if z_r is the minimizer of $(W_r(\boldsymbol{\lambda}')y_r, y_r)$ over $u_r \in S_r$ and orthogonal to the first i_r eigenvectors of $W_r(\boldsymbol{\lambda})$ then

$$(W_r(\boldsymbol{\lambda}')z_r, z_r) \leqq 0 \leqq (W_r(\boldsymbol{\lambda})z_r, z_r), \qquad 1 \leqq r < k.$$

So far, then, we have in the notation of (1)

$$\sum_{s=1}^{k} \mu_s v_{rs}(y) \leqq 0 \leqq \sum_{s=1}^{k} \mu_s v_{rs}(z), \qquad 1 \leqq r < k.$$

It follows that there is $w_r \in S_r$ (w_k being arbitrary) so that

$$\sum_{s=1}^{k} \mu_s v_{rs}(w) = 0, \qquad 1 \leqq r < k.$$

Now set

$$\sum_{s=1}^{k} \mu_s v_{ks}(w) = p$$

and eliminate $\mu_1, \cdots, \mu_{k-1}$ to obtain

$$\frac{\mu_k \det V^k(w)}{\det V^{k-1}(w)} = p.$$

From Theorem 1 and the boundedness of V_{rs} we see that p strongly decreases as a function of μ_k, and is continuous in ν. Therefore, since w_k is arbitrary in S_k, $W_k(\boldsymbol{\lambda})$ strongly decreases in λ_k (the λ_r, $1 \leqq r < k$, still being functions of λ_k) and is continuous in ν. As a result, there is $\lambda_k^{\mathbf{i}}$ so that $\rho_k^{i_k}(\boldsymbol{\lambda}^{\mathbf{i}}) = 0$, and we simply pick $x_k^{\mathbf{i}}$ as a corresponding eigenvector for $W_k(\boldsymbol{\lambda}^{\mathbf{i}})$.

This completes the existence part of the induction step. The continuous dependence part is almost identical to that for $k = 1$.

THEOREM 3. *The eigenvalues* $\boldsymbol{\lambda}^{\mathbf{i}}$ *have the property that* $\lambda_r^{\mathbf{i}}$ *increases with* i_r, $1 \leqq r \leqq k$.

Proof. That $\lambda_k^{\mathbf{i}}$ increases with i_k is a straightforward consequence of the strong decreasing behavior of $W_k(\boldsymbol{\lambda})$ in λ_k established in the proof of Theorem 2, together with the fact that ρ_k^i increases with i—see (3).

Now the construction used in Theorem 1 is symmetric in each r, so *all* nth order principal minors of $V^k(u)$ have sign $(-1)^n$. In other words, we may reorder the eigenvalue equations, putting the rth last, and repeat the arguments of Theorem 2 and the previous paragraph to obtain the desired monotonicity. This completes the proof.

The monotonicity observed in Theorem 3 may not be strict because of the possibility of multiple eigenvalues. For points $x_r \in H_r$, $y_r \in H_r$, $1 \leqq r \leqq k$, we write

$$[x, y] = \det (V_{rs}x_r, y_r).$$

We further define the *resolvent set* for the multi-parameter problem to be the set of $\boldsymbol{\lambda} \in \mathbb{R}^k$ where at least one $W_r(\boldsymbol{\lambda})$ has a bounded inverse. The *spectrum* is then the complement in $\mathbb{R}^k$ of the resolvent set. The *multiplicity* of an eigenvalue $\boldsymbol{\lambda}$ is defined as

$$\prod_{r=1}^{k} \dim \operatorname{Ker} W_r(\boldsymbol{\lambda}).$$

THEOREM 4. (i) *If* $x_1, \cdots, x_k; y_1, \cdots, y_k$ *are eigenvectors corresponding to distinct eigenvalues, then* $[x, y] = 0$.

(ii) *The spectrum consists entirely of eigenvalues.*

(iii) *Eigenvalues have finite multiplicities.*

(iv) *If* $\boldsymbol{\lambda}$ *is an eigenvalue of multiplicity* p *then there exist* p *eigenvectors* x_r^i, $1 \leqq r \leqq k$, $1 \leqq i \leqq p$ *so that* $[x^i, x^j] = \delta_{ij}$.

(v) *Eigenvalues accumulate at no finite point of* $\mathbb{R}^k$.

Proof. (i) This is an easy calculation, cf. [5, Theorem 1 or § 4].

(ii) $W_r(\boldsymbol{\lambda})$ fails to have a bounded inverse only if zero is an eigenvalue of $W_r(\boldsymbol{\lambda})$. In fact it will be the i_rth eigenvalue of $W_r(\boldsymbol{\lambda})$ for some $i_r < \infty$. This follows from the fact that $W_r(\boldsymbol{\lambda})$ has compact resolvent. Then we have $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{\mathbf{i}}$ and so the spectrum as defined above is precisely the set of eigenvalues.

(iii) When zero is an eigenvalue of $W_r(\boldsymbol{\lambda})$ it can have but finite multiplicity since $W_r(\boldsymbol{\lambda})$ has compact resolvent.

(iv) Let $\boldsymbol{\lambda}$ be an eigenvalue and consider the finite dimensional spaces $G_r = \operatorname{Ker} W_r(\boldsymbol{\lambda}) \subset H_r$. Let P_r be the orthogonal projection of H_r onto G_r and consider in the spaces G_r, the multi-parameter array of operators

$$[P_rT_rP_r \quad P_rV_{rs}P_r]_{r,s=1}^k.$$

The multi-parameter problem when posed in the spaces G_r has only one eigenvalue, namely $\boldsymbol{\lambda}$. Our earlier finite dimensional work [7] or that of Atkinson [4, Chap. 7, pp. 115–135] shows that we may find eigenvectors x satisfying the claim of this part of our theorem.

(v) The argument is easy, cf. [6, Thm. 8].

4. Geometry of the spectrum. We shall show in this section that the eigenvalues $\boldsymbol{\lambda}^{\mathbf{i}}$ form a lattice partially ordered by a certain cone $C \subset \mathbb{R}^k$. It will also be seen that when $i_r = 0$ for some r, $\boldsymbol{\lambda}^{\mathbf{i}}$ belongs to the boundary of a certain convex set contained in the resolvent set, so that in this sense, $\boldsymbol{\lambda}^{\mathbf{i}}$ is on the "outside" of the spectrum. The reader may wish to refer to § 7 for examples of the various sets constructed below. We shall use $\mathbf{a} \leqq \mathbf{b}$ for $\mathbf{a}, \mathbf{b} \in \mathbb{R}^k$ to denote $a_r \leqq b_r$, $r = 1, 2, \cdots, k$.

From now on, we shall suppress the superscript ^{k} from the notation $V^k(u)$—see (1). We introduce the set $C \subset \mathbb{R}^k$ as

$$C = \{\mathbf{a} \in \mathbb{R}^k \mid V(u)\mathbf{a} \leq \mathbf{0} \text{ for some } u_r \in S_r,\ 1 \leqq r \leqq k\}.$$

It is easily seen that if $\mathbf{a} \in C$ then $\rho\mathbf{a} \in C$, $\forall\, \rho \geqq 0$; that is C is a cone. Finally we define sets P_r^i, Z_r^i, N_r^i as the sets of points $\boldsymbol{\lambda} \in \mathbb{R}^k$ for which the ith eigenvalue of $W_r(\boldsymbol{\lambda})$ is positive, zero or negative respectively, $1 \leqq r \leqq k$, $i \geqq 0$. Obviously for each r, P_r^0 lies in the resolvent set; P_r^i, N_r^i are open and disjoint with common boundary Z_r^i.

We further claim that the set P_r^0 is convex, for if $\boldsymbol{\lambda}, \boldsymbol{\mu} \in P_r^0$ then the equation

$$W_r(\alpha\boldsymbol{\lambda} + (1-\alpha)\boldsymbol{\mu}) = \alpha W_r(\boldsymbol{\lambda}) + (1-\alpha) W_r(\boldsymbol{\mu}), \qquad 0 \leqq \alpha \leqq 1,$$

shows that in terms of the minimum eigenvalue $\rho_r^0(\boldsymbol{\lambda})$,

$$\rho_r^0(\alpha\boldsymbol{\lambda} + (1-\alpha)\boldsymbol{\mu}) \geqq \alpha\rho_r^0(\boldsymbol{\lambda}) + (1-\alpha)\rho_r^0(\boldsymbol{\mu}).$$

Thus P_r^0 is convex.

THEOREM 5. (i) *If* $\mathbf{j} \geqq \mathbf{i}$ *then* $\boldsymbol{\lambda}^{\mathbf{j}} \in \boldsymbol{\lambda}^{\mathbf{i}} + C$. *In particular, the spectrum* $\sigma \subseteq \boldsymbol{\lambda}^{\mathbf{0}} + C$.

(ii) $$\{\boldsymbol{\lambda}^{\mathbf{i}}\} = \bigcap_{r=1}^{k} Z_r^{i_r} \subseteq \bigcap_{r=1}^{k} \bar{N}_r^{i_r} \subseteq \boldsymbol{\lambda}^{\mathbf{i}} + C.$$

Proof. (i) Let $y_{rj}^{\mathbf{i}} (0 \leqq j < i_r)$ be linearly independent eigenvectors corresponding to the first i_r eigenvalues of $W_r(\boldsymbol{\lambda}^{\mathbf{i}})$. Then since 0 is the (i_r+1)th eigenvalue for $W_r(\boldsymbol{\lambda}^{\mathbf{i}})$, for any $u_r \in S_r$ such that $(u_r, y_{rj}^{\mathbf{i}}) = 0$, $0 \leqq j < i_r$, we have $(W_r(\boldsymbol{\lambda}^{\mathbf{i}})u_r, u_r) \geqq 0$. Now let u_r minimize the expression $(W_r(\boldsymbol{\lambda}^{\mathbf{j}})v_r, v_r)$ subject to $v_r \in S_r$ and $(v_r, y_{rj}^{\mathbf{i}}) = 0$, $0 \leqq j < i_r$. Since $j_r \geqq i_r$, the minimax principle yields

$$(W_r(\boldsymbol{\lambda}^{\mathbf{j}})u_r, u_r) \leqq 0 \leqq (W_r(\boldsymbol{\lambda}^{\mathbf{i}})u_r, u_r).$$

This shows that $V(u)(\boldsymbol{\lambda}^{\mathbf{j}} - \boldsymbol{\lambda}^{\mathbf{i}}) \leqq \mathbf{0}$ and so establishes the claim.

(ii) The first two relations are obvious and the third follows from an argument similar to that used for (i).

COROLLARY 1.

$$\sigma \subseteq N = \bigcap_{r=1}^{k} \bar{N}_r^0 \subseteq \boldsymbol{\lambda}^{\mathbf{0}} + C.$$

Proof. This is an easy consequence of the nesting property

$$N_r^i \subseteq N_r^{i-1}.$$

We shall discuss further properties of the set N in the next section.

When $\dim H_r = 1 + d_r < \infty$ more structure can be deduced. We state a result for the case $d_1 < \infty$.

COROLLARY 2. *Suppose* $\dim H_1 = 1 + d_1 < \infty$. *Then*

$$\sigma \subseteq N \cap \bar{P}_1^{d_1} \subseteq (\boldsymbol{\lambda}^{\mathbf{0}} + C) \cap (\boldsymbol{\lambda}^{\mathbf{a}} + C^1)$$

where $\mathbf{a}=(d_1, 0, \cdots, 0)$ *and* C^1 *is the cone defined by*

$$C^1=\{\boldsymbol{\lambda} \mid \exists\, u_r \in S_r,\ 1 \leqq r \leqq k, \text{ such that } (V(u)\boldsymbol{\lambda})_1 \geqq 0,\ (V(u)\boldsymbol{\lambda})_r \leqq 0,\ 2 \leqq r \leqq k\}.$$

Proof. Let $\boldsymbol{\lambda}^{\mathbf{i}} \in \sigma$. Then as in the proof of the theorem, we can find $u_r \in S_r$ such that

$$(W_r(\boldsymbol{\lambda}^{\mathbf{i}})u_r, u_r) \leqq 0 \leqq (W_r(\boldsymbol{\lambda}^{\mathbf{a}})u_r, u_r), \qquad 2 \leqq r \leqq k.$$

For the case $r=1$, note that $W_1(\boldsymbol{\lambda}^{\mathbf{a}})$ has 0 as its last eigenvalue so that $(W_1(\boldsymbol{\lambda}^{\mathbf{a}})u_1, u_1) \leqq 0$, $\forall\, u_1 \in S_1$. In particular, we take u_1 to be the eigenvector corresponding to the zero eigenvalue of $W_1(\boldsymbol{\lambda}^{\mathbf{i}})$. Then we obtain

$$(W_1(\boldsymbol{\lambda}^{\mathbf{i}})u_1, u_1)=0 \geqq (W_1(\boldsymbol{\lambda}^{\mathbf{a}})u_1, u_1).$$

This shows that $\boldsymbol{\lambda}^{\mathbf{i}}-\boldsymbol{\lambda}^{\mathbf{a}} \in C^1$. A like argument establishes that if $\boldsymbol{\lambda} \in \bar{P}_1^{d_1}$ then $\boldsymbol{\lambda} \in \boldsymbol{\lambda}^{\mathbf{a}}+C^1$. The inclusion $\sigma \subset \bar{P}_1^{d_1}$ is obvious.

If $d_r<\infty$ for all r, then 2^k cones can be produced in this manner. For each "extreme" index $\mathbf{i}$ where i_r is either 0 or d_r we generate a cone as in Corollary 2. We then have the obvious generalizations of Corollary 2.

5. Vectorial ranges. In [7] we defined and investigated a vectorial range for multi-parameter eigenvalue problems. In this section we present the corresponding theory for our Hilbert space setting.

Let $H=\otimes_{r=1}^k H_r$ denote the tensor product of the spaces H_r. $D \subset H$ will denote decomposable tensors $u=u_1 \otimes \cdots \otimes u_k$ where each $u_r \in \mathscr{D}(T_r)$, $1 \leqq r \leqq k$. For such a point u we may define $T_r^\dagger u=u_1 \otimes \cdots \otimes u_{r-1} \otimes T_r u_r \otimes u_{r+1} \otimes \cdots \otimes u_k$ and similarly for $V_{rs}^\dagger$. We can then extend $T_r^\dagger$ by linearity and in the case of $V_{rs}^\dagger$ by linearity and continuity. This introduces the possibility of defining operators $\Delta_0, \cdots, \Delta_k$ in H by means of the formal determinantal expansion

$$\sum_{s=0}^k \alpha_s \Delta_S=\det\begin{vmatrix} \alpha_0 & \alpha_1 & \cdots & \alpha_k \\ T_1^\dagger & V_{11}^\dagger & \cdots & V_{1k}^\dagger \\ \cdot & \cdot & \cdot & \cdot \\ T_{1k}^\dagger & V_{k1}^\dagger & \cdots & V_{kk}^\dagger \end{vmatrix}.$$

Δ_0 may be defined on all of H while we have $\Delta_1, \cdots, \Delta_k$ defined on the algebraic tensor product of the domains $\mathscr{D}(T_r)$, $1 \leqq r \leqq k$, which we now denote by E. In line with our earlier notation we write $[u, v]=(\Delta_0 u, v)$ for $u, v \in H$ and $S=\{u \in H \mid [u, u]=1\}$.

The *vectorial range* for our multi-parameter problem may now be defined as the set of vectors in $\mathbb{R}^k$ with rth component $\delta_r=(\Delta_r u, u)$ for $u \in S \cap E$, such a vector being denoted by $\boldsymbol{\delta}(u)$. In the finite dimensional setting [7] we showed that $\boldsymbol{\delta}(S \cap E)$ is the closed convex hull of σ. Here we shall examine the sets $\boldsymbol{\delta}(S \cap D)$ and $\boldsymbol{\delta}(S \cap F)$ where F is the space of all finite linear combinations of eigenvectors $\otimes_{r=1}^k u_r^{\mathbf{i}}$.

THEOREM 6. *If each space* H_r *is of infinite dimension then the decomposable vectorial range* $\boldsymbol{\delta}(S \cap D)$ *is the set* N *of Theorem* 5, *Corollary* 1 *and is thus the complement of the union of* k *open convex sets in* $\mathbb{R}^k$.

Proof. By virtue of Cramer's rule we see that $\boldsymbol{\lambda} \in \boldsymbol{\delta}(S \cap D)$ if and only if there exists $u_r \in S_r$ such that

$$(T_r u_r, u_r) + \sum_{s=1}^{k} \lambda_s (V_{rs} u_r, u_r) = 0, \qquad 1 \leqq r \leqq k.$$

This is equivalent to the statement $\boldsymbol{\lambda} \in N$—recall that in the case $\dim H_r = \infty$, $W_r(\boldsymbol{\lambda})$ has eigenvalues accumulating only at $+\infty$.

THEOREM 7. *If* $\dim H_r = 1 + d_r < \infty$ *for each* r *then* $\boldsymbol{\delta}(S \cap D)$ *is the complement of* $2k$ *open convex sets in* $\mathbb{R}^k$. *Explicitly*

$$\boldsymbol{\delta}(S \cap D) = \bigcap_{r=1}^{k} (\bar{N}_r^0 \cap \bar{P}_r^{d_r}) = N \cap P.$$

Proof. The relation $\boldsymbol{\delta}(S \cap D) = \bigcap_{r=1}^{k} (\bar{N}_r^0 \cap \bar{P}_r^{d_r})$ is established as in Theorem 6. Further we have

$$\bigcap_{r=1}^{k} (\bar{N}_r^0 \cap \bar{P}_r^{d_r}) = \sim \left[\bigcup_{r=1}^{k} (P_r^0 \cap N_r^{d_r}) \right].$$

We showed earlier that P_r^0 is convex. A similar argument establishes the convexity of $N_r^{d_r}$.

We turn now to nondecomposable tensors.

THEOREM 8. *The eigen-vectorial range* $\boldsymbol{\delta}(S \cap F)$ *is the convex hull* co σ *of the spectrum.*

Proof. If $\boldsymbol{\lambda} \in \text{co}\, \sigma$ then by Carathéodory's theorem we can find $k+1$ eigenvalues $\boldsymbol{\lambda}^{\mathbf{i}_j}$ and nonnegative constants α_j, $0 \leqq j \leqq k$, so that $\sum_{j=0}^{k} \alpha_j = 1$ and $\boldsymbol{\lambda} = \sum_{j=1}^{k} \alpha_j \boldsymbol{\lambda}^{\mathbf{i}_j}$. Then setting $u = \sum_{j=0}^{k} \alpha_j^{1/2} u^{\mathbf{i}_j}$ (where the eigenvectors $u^{\mathbf{i}_j}$ are normalized so as to belong to S and are $[\cdot,\cdot]$ orthogonal) we obtain $[u, u] = 1$ and $\lambda_r = (\Delta_r u, u)$ so that $\boldsymbol{\lambda} \in \boldsymbol{\delta}(S \cap F)$.

Conversely, if $u \in H$ is expressible as a finite sum $u = \sum \beta_j u^{\mathbf{i}_j}$ then $u \in S$ if and only if $\sum |\beta_j|^2 = 1$ and further we have

$$(\Delta_r u, u) = \sum |\beta_j|^2 \lambda_r^{\mathbf{i}_j}$$

whence $\boldsymbol{\lambda} \in \text{co}\, \sigma$.

In the case $\dim H_r < \infty$, we have $F = H$ so that our last result is sufficient to characterize the vectorial range and we obtain $\sigma \subset N \cap P \subset \text{co}\, \sigma$, $\text{co}\,(N \cap P) = \text{co}\, \sigma$. In infinite dimensions some caution should be exercised since it is not clear that H is a Hilbert space under $[\cdot,\cdot]$ as inner product. This situation can be achieved by assuming $\Delta_0 \gg 0$ on H, in which case it is known that $\bar{F} = H$ [6, Thm. 10].

6. Parametric dependence. In this section we shall allow the operators in our multi-parameter problem to depend on a parameter ν, which for simplicity will be taken from a Euclidean space. For our first result the V_{rs} and the domains $\mathscr{D}(T_r)$ are to be independent of ν.

THEOREM 9. *Let* T_r *increase with* ν; *that is for* $\nu' \geqq \nu$ *we have* $T_r(\nu') - T_r(\nu)$ *is positive semi-definite on* $\mathscr{D}(T_r)$. *Then* $\lambda_r^{\mathbf{i}}(\nu') \in \lambda_r^{\mathbf{i}}(\nu) + C$ *for each* $\mathbf{i}$ *and* r.

Proof. Let $\nu' \geqq \nu$, and set $\lambda_r^{\mathbf{i}}(\nu) = \lambda_r$, $\lambda_r^{\mathbf{i}}(\nu') = \lambda_r'$, $W_r(\lambda^{\mathbf{i}}(\nu), \nu) = W_r$ etc. Now let y_{rj}, $0 \leqq j < i_r$ be linearly independent eigenvectors of W_r and let u_r minimize $(W_r' v_r, v_r)$ subject to $v_r \in S_r$ and $(v_r, y_{rj}) = 0$, $0 \leqq j < i_r$ to give

$$(W_r' u_r, u_r) \leqq 0 \leqq (W_r u_r, u_r).$$

Then

$$((T_r(\nu')-T_r(\nu))u_r, u_r)+(V(u)(\boldsymbol{\lambda}'-\boldsymbol{\lambda}))_r \leqq 0$$

so nonnegativity of the first term gives the desired conclusion.

For the next result we allow the V_{rs} to vary with ν, but the T_r are assumed independent of ν.

THEOREM 10. *Let the operators V_{rs} depend continuously on ν. Then so does the spectrum, in the following sense. If as $\nu_j \to \nu$, a sequence of eigenvalues $\boldsymbol{\lambda}(\nu_j)$, (not necessarily of fixed index) accumulates at $\boldsymbol{\lambda}$, then $\boldsymbol{\lambda}$ is an eigenvalue of the limiting problem. Conversely, every eigenvalue for ν is the limit of some sequence of eigenvalues of ν_j. Further if $u(\nu_j)=\otimes_{r=1}^k u_r(\nu_j)$ is a corresponding sequence of eigenvectors with $\|u_r(\nu_j)\|=1$, then each sequence $u_r(\nu_j)$ has a limit u_r and $u=\otimes_{r=1}^k u_r$ is an eigenvector of the limiting problem corresponding to $\boldsymbol{\lambda}$.*

Proof. Continuity of $\boldsymbol{\lambda}^{\mathbf{i}}$ follows from the arguments used for the inductive step in Theorem 2. Conversely, any $\boldsymbol{\lambda}$ for the limiting problem has an index, say $\boldsymbol{\lambda}=\boldsymbol{\lambda}^{\mathbf{i}}(\nu)$. Then $\boldsymbol{\lambda}^{\mathbf{i}}(\nu_j)$, $j=1,2,\cdots$ is a suitable sequence.

For the final part of the theorem we have for any nonreal number θ

$$(T_r+\theta I)u_r(\nu_j)+\sum_{s=1}^k \lambda_s(\nu_j)V_{rs}(\nu_j)u_r(\nu_j)=\theta u_r(\nu_j),$$

$$u_r(\nu_j)+\sum_{s=1}^k \lambda_s(\nu_j)(T_r+\theta I)^{-1}V_{rs}(\nu_j)u_r(\nu_j)=\theta(T_r+\theta I)^{-1}u_r(\nu_j). \tag{5}$$

Since $\|u_r(\nu_j)\|=1$ we can assume (by taking a subsequence if necessary) that $u_r(\nu_j)$ has a weak limit u_r. We now appeal to the compactness of $(T_r+\theta I)^{-1}$ and the continuous dependence of V_{rs} upon ν. From (5) we see that $u_r(\nu_j)$ converges in norm to u_r, so that we have, from (5),

$$u_r+\sum_{s=1}^k \lambda_s(T_r+\theta I)^{-1}V_{rs}(\nu)u_r=\theta(T_r+\theta I)^{-1}u_r.$$

We may now claim that $u_r\in\mathscr{D}(T_r)$ and so obtain

$$T_r u_r+\sum_{s=1}^k \lambda_s V_{rs}(\nu)u_r=0$$

to establish the result.

Our final result of this section concerns the situation in which each V_{rs} is allowed to depend in a Lipschitz manner on a parameter ν.

THEOREM 11. *Let ν range over a compact set and let γ—the constant of the definiteness condition (2)—be independent of ν. If the operators V_{rs} are Lipschitz in ν then so are the eigenvalues $\boldsymbol{\lambda}^{\mathbf{i}}$.*

Proof. The transformation of Theorem 1 now provides determinant bounds which are independent of ν. As a result the bound β in (4) remains independent of ν and the subsequent argument easily extends to show that $\lambda^i(\nu)$ is Lipschitz in ν for $k=1$. The inductive step in Theorem 2 generalizes this to the vector eigenvalue $\boldsymbol{\lambda}^{\mathbf{i}}(\nu)$.

7. Examples. We present here two examples designed to illustrate the geometric properties of the spectrum discussed in earlier sections.

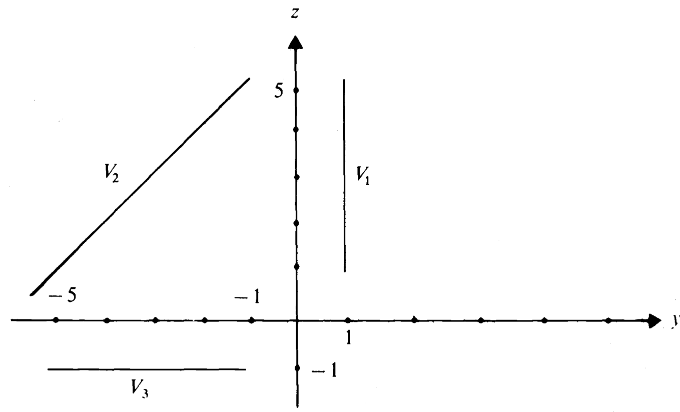

FIG. 1. *Cross sections of* V_1, V_2, V_3 *at* $x = -5$.

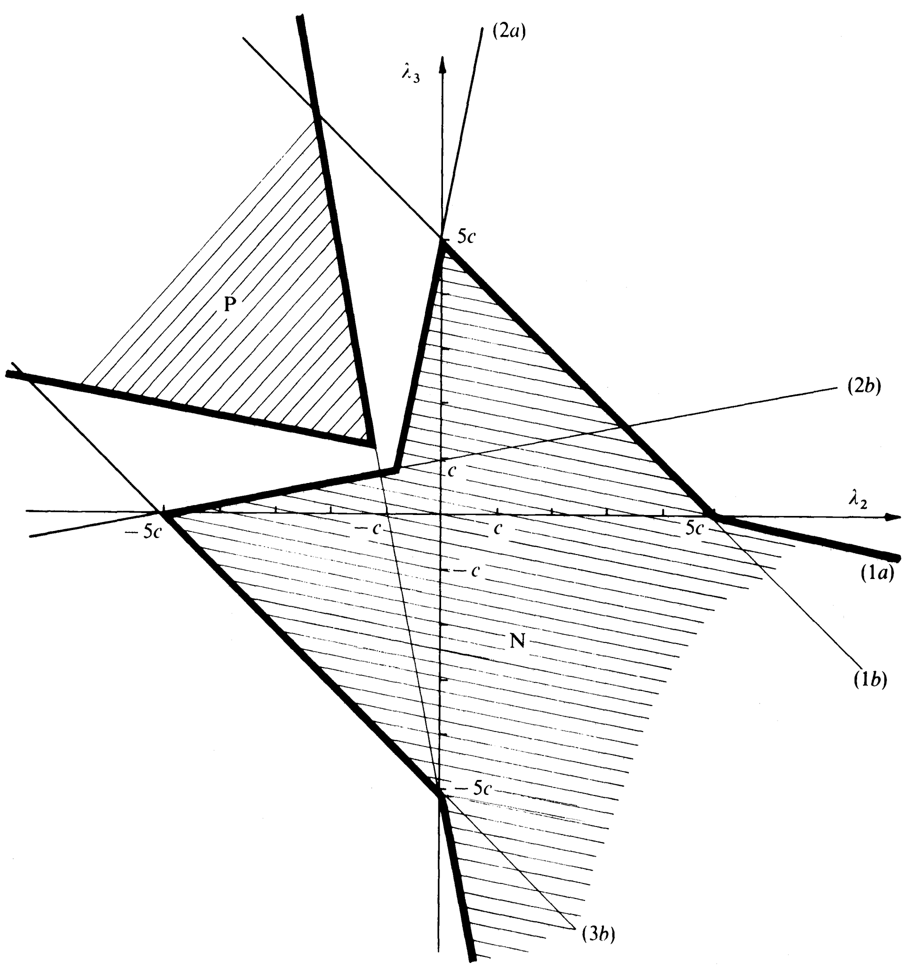

FIG. 2. *Cross sections of P and N at* $\lambda_1 = c > 0$.

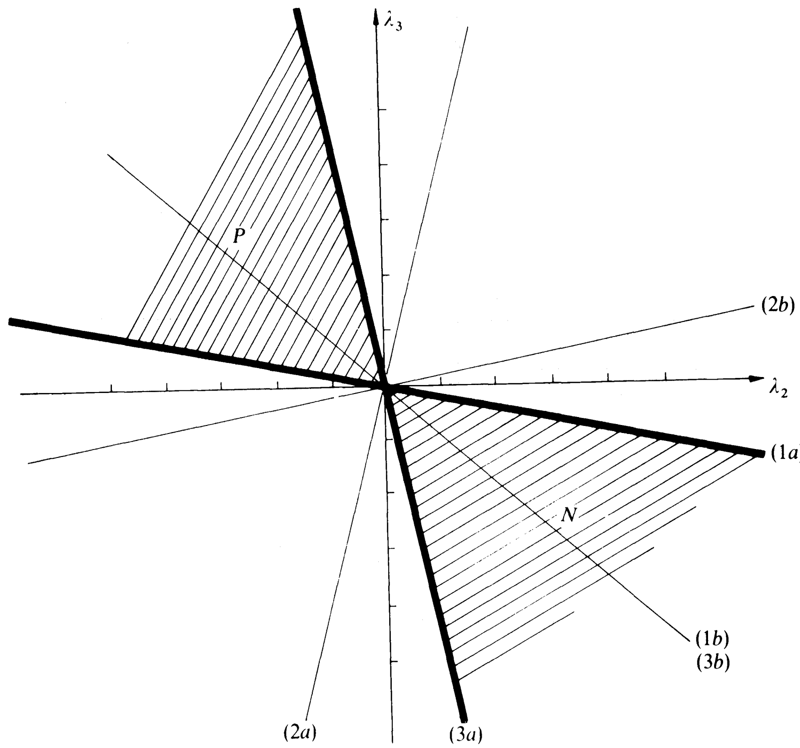

FIG. 3. *Cross sections of N and P at* $\lambda_1 = 0$.

Example 1. Take $H_1 = H_2 = H_3 = \mathbb{C}^2$ with operators T_r, V_{rs} given by the array

$$\begin{bmatrix} 0 & -5I & I & \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix} \\ 0 & -5I & \begin{bmatrix} -5 & 0 \\ 0 & -1 \end{bmatrix} & \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix} \\ 0 & -5I & \begin{bmatrix} -5 & 0 \\ 5 & -1 \end{bmatrix} & -I \end{bmatrix}$$

Notice that all principal minors of order n do have sign $(-1)^n$, so no extra transformation (per Theorem 1) is needed. This is obvious by inspection for $n = 1$ and 2, while for $n = 3$ we observe that $u_1 = u_2 = u_3 = (0, 1)$ gives $\det V(u) = -5$. To show that $\det V(u)$ has constant sign for all u_r we examine the cones V_1, V_2 and V_3 generated by

$$V_r = \{(V_{rs}u_r, u_r) \mid u_r \in H_r\} \subset \mathbb{R}^3.$$

Using points $u_r \in H_r$ of unit norm we see that the points in V_r are of the form $(-5, y, z)$ as displayed in Fig. 1.

It is now geometrically clear that the conditions of [4, Thm. 9.6.1, p. 152] hold so that we can claim the definiteness condition to be satisfied. Alternatively we could have calculated a rather cumbersome 3×3 determinant.

With each $T_r = 0$, it is obvious that the spectrum for this problem consists of only one eigenvalue, zero, which has multiplicity eight.

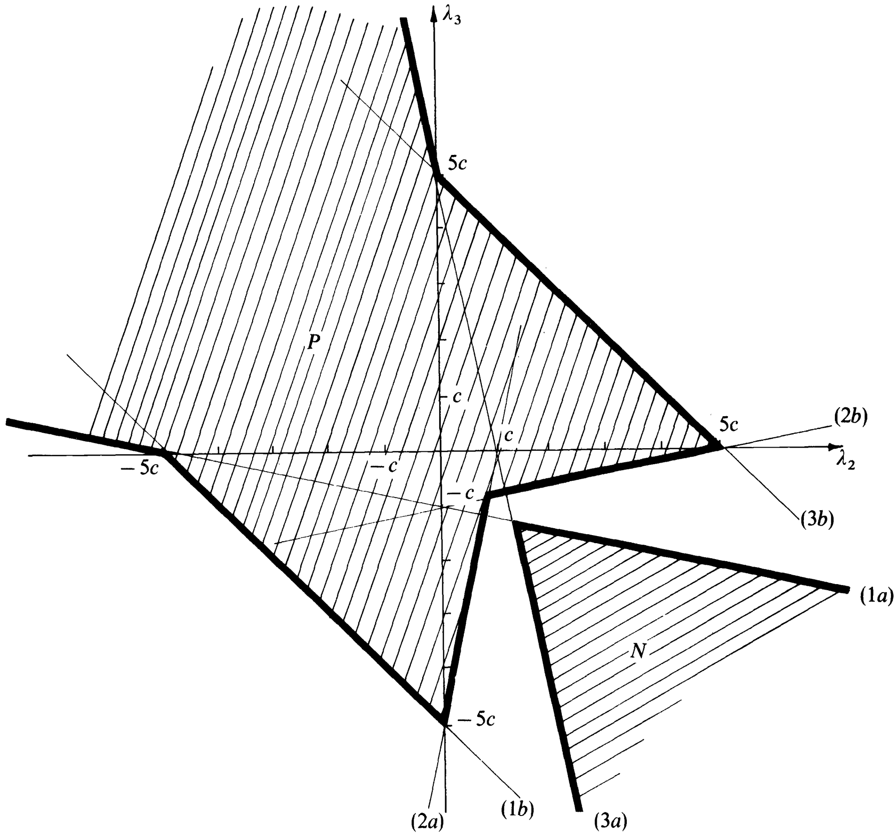

FIG. 4. *Cross-sections of N and P at* $\lambda_1 = c < 0$.

For $\boldsymbol{\lambda} \in \mathbb{R}^3$ the operators $W_r(\boldsymbol{\lambda})$ have eigenvalues as follows:

$$\begin{array}{llll} W_1(\boldsymbol{\lambda}): & -5\lambda_1+\lambda_2+5\lambda_3 \quad (1a), & -5\lambda_1+\lambda_2+\lambda_3 & (1b); \\ W_2(\boldsymbol{\lambda}): & -5\lambda_1-5\lambda_2+\lambda_3 \quad (2a), & -5\lambda_1-\lambda_2+5\lambda_3 & (2b); \\ W_3(\boldsymbol{\lambda}): & -5\lambda_1-5\lambda_2-\lambda_3 \quad (3a), & -5\lambda_1-\lambda_2-\lambda_3 & (3b). \end{array}$$

The labels (1a), · · · , (3b) are cross-referenced in Figs. 2–4 which display cross sections of the sets N and P for $\lambda_1 > 0$, $\lambda_1 = 0$ and $\lambda_1 < 0$. Note that neither N nor P is convex. In this case $N \cap P$ coincides with the spectrum. We could, by slight perturbations of the T_r, split the multiple eigenvalue **0** into eight simple eigenvalues.

Example 2. Let $H_1 = L^2[0, \pi/2]$ (Lebesgue measure) with T_1 given by

$$\mathscr{D}(T_1) = \{f \in L^2[0, \pi/2] | df/dx \text{ is absolutely continuous}, \\ (d^2f/dx^2) - f \in L^2[0, \pi/2], f(0) = f'(\pi/2) = 0\},$$

$$\text{for } f \in \mathscr{D}(T_1), \quad T_1 f = \frac{d^2 f}{dx^2} - f.$$

Let $H_2 = \mathbb{C}^2$ and consider the multiparameter array

$$\begin{bmatrix} T_1 & -I & I \\ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix} & \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \end{bmatrix}.$$

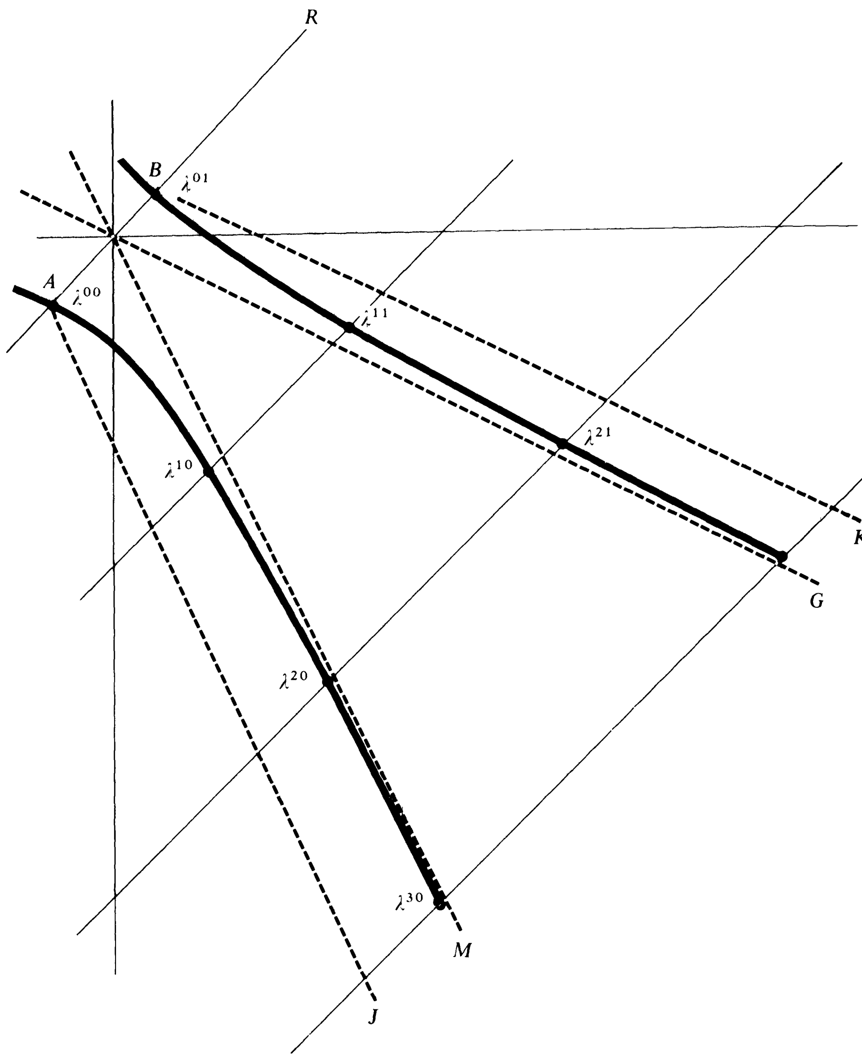

FIG. 5. *Spectral diagram for Example* 2.

Note that V_{11}, $V_{22} \ll 0$ and (2) is satisfied, so again all principal minors of order n have sign $(-1)^n$.

Solving $W_1(\boldsymbol{\lambda})f = 0$ we obtain $\lambda_1 - \lambda_2 + 1 = 1, 3, 5, \cdots$, so that $\lambda_1 - \lambda_2 = 0, 2, 4, \cdots$. Solving $W_2(\boldsymbol{\lambda})x = 0$ we obtain $(2\lambda_1 + \lambda_2)(\lambda_1 + 2\lambda_2) = 1$ as the eigenvalue condition. Eigenvalues occur then at the intersections of the straight lines $\lambda_1 - \lambda_2 = 2n$, $n = 0, 1, \cdots$ and the hyperbola above. These are marked on Fig. 5. We make the following identifications:

$$N_1^0 = \{\boldsymbol{\lambda} | \lambda_1 < \lambda_2\}, \quad Z_1^0 = \{\boldsymbol{\lambda} | \lambda_1 = \lambda_2\}, \quad P_1^0 = \{\boldsymbol{\lambda} | \lambda_1 > \lambda_2\}.$$

N_2^0 can be obtained by solving the inequalities

$$[(2\lambda_1 + \lambda_2)(\lambda_1 + 2\lambda_2) < 1]$$

or

$$[(2\lambda_1 + \lambda_2)(\lambda_1 + 2\lambda_2) < 1 \text{ and } \lambda_1 + \lambda_2 > 0].$$

This gives the region above the lower branch of the hyperbola. These calculations are elementary exercises and lead to the conclusion that N is closed unbounded nonconvex region bounded by RAM, while $\boldsymbol{\lambda}^0 + C$ is the set bounded by RAJ. We also see that P_2^1 is the set below the upper branch of the hyperbola so that

$$\boldsymbol{\delta}(S \cap D) = GBAM = N \cap P_2^1.$$

We have labeled the eigenvalues so as to display the monotonicity in the indices as discussed in Theorem 3. Finally we note that co σ is not closed in this example. In fact this set will be given by the convex region enclosed by $JABK$, the open boundaries AJ and BK not being included.

REFERENCES

[1] F. M. ARSCOTT, *Periodic Differential Equations, an Introduction to Mathieu, Lamé and Allied Functions*, Macmillan, New York, 1964.

[2] ———, *Two-parameter eigenvalue problems in differential equations*, Proc. London Math. Soc., 14 (1964), pp. 459–470.

[3] F. V. ATKINSON, *Multiparameter spectral theory*, Bull. Amer. Math. Soc., 74 (1968), pp. 1–27.

[4] ———, *Multiparameter Eigenvalue Problems, Vol. I: Matrices and Compact Operators*, Academic Press, New York, 1972.

[5] P. J. BROWNE, *A multiparameter eigenvalue problem*, J. Math. Anal. Appl., 38 (1972), pp. 553–568.

[6] ———, *Abstract multiparameter theory*, Ibid., 60 (1977), pp. 259–273.

[7] P. A. BINDING AND P. J. BROWNE, *A variational approach to multiparameter eigenvalue problems for matrices*, this Journal, 8 (1977), pp. 763–777.

[8] A. WEINSTEIN AND W. STENGER, *Methods of Intermediate Problems for Eigenvalues*, Academic Press, New York, 1972.

SIAM J. MATH. ANAL.
Vol. 9, No. 6, December 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0006 $01.00/0

REPRESENTATION OF SPECIAL FUNCTIONS BY DIFFERINTEGRAL AND HYPERDIFFERENTIAL OPERATORS*

DO TAN SI†

Abstract. The use of differintegral operators D^ν with $D=d/dz$ and $\nu\in\mathbb{C}$ and hyperdifferential operators, together with some operational calculus manipulations leads to new formulae of representation of usual special functions. These formulae are useful for obtaining generating functions and differential recurrence relations of orthogonal polynomials.

1. Introduction. The main purpose of this work is first to put the hypergeometric functions into concise form with differintegral operators D^ν where $D\equiv d/dz$ and $\nu\in\mathbb{C}$ [9]. Second, it is to show that usual orthogonal polynomials are related to monomials $C_n z^n$ in some transformations realizable by hyperdifferential operators. The latter are defined as differential operators of infinite order with variable coefficients:

$$F(z,D)\equiv\sum_{i,j=0}^{\infty} a_{ij} z^i D^j. \tag{1.1}$$

To this end, we discuss—in § 2— the mathematical background of the method used in this work. This method consists chiefly in the definition of the operator D^ν in such a way that it may be used to obtain the most general solution of a linear differential equation. In this context, the equation $D^n y=0$, for example, leads to $y=D^{-n}0=F(D)z^n/n!$ with arbitrary analytic function $F(z)$ satisfying $F(0)=0$. On the other hand, we use hereafter hyperdifferential operators without discussing further its theory and algebra. To our knowledge, one may find a rigorous definition of hyperdifferential operators in the works of Treves [12], Miller and Steinberg [8], etc. Besides, the algebra developed about the Baker–Campbell–Hausdorff relation seems well covered in the work of Wilcox [13]. In the above works, one finds many applications of hyperdifferential calculus either in quantum mechanical problems [8], [13] or in the integration of partial differential equations [13], etc. We note also the recent work of Wolf [14], [15] where canonical and integral transforms are realized through hyperdifferential operators.

In § 3, we derive the results on hypergeometric functions and orthogonal polynomials, using operators D^ν and hyperdifferential ones. The formulae obtained are very suitable for practical calculations of these special functions. They are also useful in the derivation of differential recurrence relations and generating functions of orthogonal polynomials.

2. Background of the method.

2.1. In dealing with differintegral operators or fractional integrals we start from the following definition of Riemann [9]:

$$D^\mu z^\nu=\frac{\Gamma(\nu+1)}{\Gamma(\nu-\mu+1)}z^{\nu-\mu};\qquad \nu,\mu\in\mathbb{C}. \tag{2.1}$$

We note that many other definitions are known [3], [4], [9] and perhaps are more suitable for other problems.

* Received by the editors December 11, 1975, and in final revised form May 2, 1977.
† Faculté des Sciences, Université de l'Etat à Mons, B-7000 Mons, Belgium.

From (2.1) one derives the equation

$$D^\nu z^\nu/\Gamma(\nu+1)=1. \tag{2.2}$$

If now $F(z)$ is an arbitrary analytic function, then applying the operator $F(D)$ on both sides of (2.2), one gets:

$$D^\nu F(D)z^\nu/\Gamma(\nu+1)=F(D)\cdot 1=c$$

so that

$$D^{-\nu}c=F(D)z^\nu/\Gamma(\nu+1); \qquad c=F(0). \tag{2.3}$$

In the current literature, one usually chooses $F(z)=c$ in order to insure the linearity of the operator D^ν and the unicity of the inverse operator $D^{-\nu}$. Nevertheless we think that in a problem where these conditions may be omitted, one may assume $F(z)$ to be an arbitrary function with $F(0)=c$. Only at the end will $F(z)$ be determined from subsidiary conditions, as in the case of a differential equation for which a solution corresponding to some boundary conditions is deduced from the general one.

In conclusion, we propose to replace the definition (2.1) by:

$$D^\mu z^\nu=D^\mu(z^\nu+0)=\frac{\Gamma(\nu+1)}{\Gamma(\nu-\mu+1)}z^{\nu-\mu}+\frac{F(D)z^{-\mu}}{\Gamma(1-\mu)} \tag{2.4}$$

with $F(D)=0$ if we want D^μ to be linear, or only $F(0)=0$ if we do not need the linearity of D^μ in a closed loop of the calculations.

The determination of $F(z)$ in each problem is essentially a matter of identification of coefficients. For example, if a set of polynomials $P_n(z)$ may be put into the form

$$P_n(z)=F(D)z^n=\sum_{m=0}^{\infty} f_m D^m z^n \tag{2.5}$$

then

$$P_n(0)=n!f_n \tag{2.6}$$

and

$$F(D)=\sum_{n=0}^{\infty} P_n(0)D^n/n! \tag{2.7}$$

It is well known that functions $f(z)$ which are linear combinations of exponentials have the property [13]

$$[z,f(D)]=-\frac{\partial}{\partial D}f(D)=-f'(D). \tag{2.8}$$

In the case where $f(D)=D^\nu$, using (2.4) and the fact that

$$(z-1)D^\nu 0=-\nu D^{\nu-1}0=-\nu F(D)z^{1-\nu}/\Gamma(2-\nu) \tag{2.9}$$

one gets the equation

$$[z,D^\nu]z^\mu=-\frac{\nu\Gamma(1+\mu)}{\Gamma(2+\mu-\nu)}z^{\mu-\nu+1}+(z-1)D^\nu 0=-\nu D^{\nu-1}z^\mu \tag{2.10}$$

which shows that (2.8) still holds for D^ν.

2.2. Also, when dealing with the operational calculus in this work, we often use the following identity, easy to prove:

$$a(z)D+b(z)=a(z)\,e^{-u(z)}D\,e^{u(z)} \tag{2.11}$$

where $u(z)$ is a primitive of the function $b(z)/a(z)$. Equation (2.11) leads directly to the inversion formula

$$[a(z)D+b(z)]^{-n}=e^{-u(z)}D^{-n}\,e^{u(z)}a^{-1}(z) \tag{2.12}$$

and the following:

$$a(z)f(D)a^{-1}(z)=f(D+\ln' a(z)),\quad \forall f \text{ analytic} \tag{2.13}$$

which are very useful in the studies of orthogonal polynomials.

3. Contribution to the studies of special functions.

3.1. Hypergeometric, confluent hypergeometric functions. First, we consider the differential equation:

$$Ay\equiv[(mz^2+nz+p)D^2+(qz+r)D+s]y=0 \tag{3.1}$$

which was resolved by Liouville [6], [9] and Holmgren [5] more than a century ago, utilizing fractional integrals. The main part of the method of Liouville may be described in a concise manner as follows.

We have from (2.8) the relations

$$D^{-\nu}zD^{\nu}=z-\nu D^{-1}, \tag{3.2}$$

$$D^{-\nu}z^2D^{\nu}=z^2-2\nu zD^{-1}+(\nu^2+\nu)D^{-2} \tag{3.3}$$

which lead to:

$$\begin{aligned} D^{-\nu}AD^{\nu}&=(mz^2+nz+p)D^2+(qz+r-2m\nu z-n\nu)D\\ &\quad+(m\nu^2+(m-q)\nu+s)\equiv I(z)D^2+J(z)D+K. \end{aligned} \tag{3.4}$$

Now, if the parameter ν satisfies the equation

$$K\equiv m\nu^2+(m-q)\nu+s=0 \tag{3.5}$$

one gets from (3.1) and (3.4) the first order linear differential equation in $D^{1-\nu}y$:

$$D^{-\nu}AD^{\nu}D^{-\nu}y=(ID+J)D^{1-\nu}y=D^{-\nu}0. \tag{3.6}$$

Using (2.3) and (2.12) one can then write

$$y=D^{\nu-1}\,e^{-u(z)}\int e^{u(z)}(mz^2+mz+p)^{-1}F(D)z^{\nu}/\Gamma(1+\nu) \tag{3.7}$$

where $F(0)=0$ and $u(z)=\int J(z)I^{-1}(z)$.

Putting now $y_{-1}=D^{\nu-1}0$ and labeling the solutions (3.7) corresponding to $F(D)=0$ and $F(D)=D^k$, $k>0$, by y_0 and y_k respectively, we see that:

(i) if $\nu=0,\,-1,\,-2,\cdots$, then y_{-1} and y_0 are two independent solutions, all the y_k being equal to y_0;

(ii) if $\nu=1,2,\cdots$, then $y_{-1}=0$, $y_0=y_{\nu+1}=y_{\nu+2}=\cdots$ but $y_0\neq y_k$ with $k\leqq\nu$ so that one has at least two solutions;

(iii) if $\nu\neq 0,\,\pm1,\,\pm2,\cdots$, then y_{-1} does not satisfy (3.6) nor a fortiori (3.1) but $y_0\neq y_k$, $\forall k$.

It then remains for us to justify that whenever $y_k \neq y_0$, they are proportional to y_1; i.e. for any value of ν and any choice of $F(D)$, one gets two and only two independent solutions for (3.1), as it must be.

To this end, we remark that, when applied on z^ν, D^k with $k>1$ is proportional to zD^{k+1} or z^2D^{k+2}. Thus there exists a linear combination (l.c.) of D^k and D^{k+1} such that l.c. $(D^k, D^{k+1}) = J(z)D_{k+1}$. Noting that $u'(z)=I^{-1}(z)J(z)$, the same l.c. between y_k and y_{k+1} gives rise to

$$\text{l.c. } (y_k, y_{k+1}) = D^{\nu-1}e^{-u}\int e^u I^{-1} J D^{k+1} z^\nu/\Gamma(1+\nu) \tag{3.8}$$

$$= D^{\nu+k} z^\nu/\Gamma(1+\nu) - D^{\nu-1}e^{-u}\int e^u D^{k+1} z^\nu/\Gamma(1+\nu) \tag{3.9}$$

$$= -D^{\nu-1}e^{-u}\int e^u I^{-1}(mz^2+nz+p)D^{k+1}z^\nu/\Gamma(1+\nu). \tag{3.10}$$

As in the equation (3.10), z^2D^{k+1} is proportional to D^{k-1} and zD^{k+1} to D^k, its right hand side is a l.c. (y_{k-1}, y_k, y_{k+1}) if $p \neq 0$ or a l.c. (y_{k-1}, y_k) if $p=0$. The latter case may always be realized by a suitable change of variable. Thus the proof of the proportionality between the y_k's which differ from y_0 is found.

Second, for the concrete case of hypergeometric equations:

$$z(z-1)y'' + ((a+b+1)z-c)y' + aby = 0, \tag{3.11}$$

we get $\nu = a$ (or b) and $u(z) = \ln(1-z)^{b+1-c}z^{c-a}$.

This leads to the result that hypergeometric functions may be put into the general form:

$${}_2F_1(a,b;c;z) = CD^{a-1}(1-z^{c-b-1}z^{a-c}\int (1-z)^{b-c}z^{c-k-1}/\Gamma(a-k+1) \tag{3.12}$$

where k is a suitable positive integer (giving y_k) and where the constant C and the constant of integration (giving y_0) are both determined by the properties ${}_2F_1(a,b;c;0)=1$, ${}_2F_1'(a,b;c;0)=ab/c$. Similarly, for confluent hypergeometric functions we get

$${}_1F_1(a;b;z) = CD^{a-1}e^z z^{a-b}\int e^{-z}z^{b-k-1}/\Gamma(a-k+1). \tag{3.13}$$

3.2. Hermite polynomials. The representation of Hermite polynomials by hyperdifferential operators derived by Wolf [14] may also be obtained in the Bargman space [2] or in any other space provided the operators z and D in this space are Hermitian conjugates:

$$z^+ = D. \tag{3.14}$$

In fact, from (3.14) and (2.11) one has successively:

$$-(D-z) = (D-z)^+, \tag{3.15}$$

$$-e^{z^2/2}De^{-z^2/2} = e^{-D^2/2}z\,e^{D^2/2},$$

$$e^{z^2/2}(-D)^n e^{-z^2/2} = e^{-D^2/2}z^n e^{D^2/2}. \tag{3.16}$$

Hermite polynomials, defined by the Rodrigues formula, arise from the application of both sides of (3.16) on a constant, followed by a change of variable $z \to z\sqrt{2}$:

$$H_n(z) = e^{-D^2/4}(2z)^n. \tag{3.17}$$

Furthermore, since the operator in the left hand side of (3.13) is nothing but $(z-D)^n = 2^{n/2}a^{+n}$ where a^+ describes the creation operator of a harmonic oscillator quantum, we have

$$H_n(z/\sqrt{2}) = a^{+n}\cdot 1 \tag{3.18}$$

and

$$(a^+)^n e^{-2^2/2} = 2^{-n/2} e^{-z^2/2} H_n(z). \tag{3.19}$$

From the formula (3.17), one gets the general generating function

$$g_k(z,t) = \sum_{n=0}^{\infty} a_n H_{n+k}(z)t^n = e^{-D^2/4}(2z)^k f(2zt) \tag{3.20}$$

where obviously

$$f(z) = \sum_{n=0}^{\infty} a_n z^n. \tag{3.21}$$

Using (3.16), we can also write

$$g_k(z,t) = (-)^k e^{z^2} D^k e^{-z^2} e^{-D^2/4} f(2zt). \tag{3.22}$$

The calculation of the transform by $\exp(-D^2/4)$ of $f(2zt)$ is particularly simple when $f(z)$ is an eigenfunction of the operator D^2. When $f(z)$ is a Gaussian, its transform by $\exp(-D^2/4)$ may also be easily calculated. As e^{az^2} satisfies the differential equation $(D-2az)y=0$, its transform by e^{bD^2} must satisfy the equation

$$e^{bD^2}(D-2az)e^{-bD^2}y = 0, \quad \text{i.e. } \{(1-4ab)D - 2az\}y = 0.$$

That implies the relation

$$e^{bD^2} e^{az^2} = (1-4ab)^{-1/2} e^{a(1-4ab)^{-1}z^2}; \qquad 4az < 1. \tag{3.23}$$

Equations (3.23) give rise to the bilinear generating function

$$\begin{aligned} \sum_{n=0}^{\infty} H_n(x)H_{n+k}(y)t^n/n! &= e^{-D_y^2/4 - D_x^2/4}(2y)^k e^{4xyt} \\ &= e^{-D_y^2/4}(2y)^k e^{-(2yt)^2+4xyt} \\ &= (1-4t^2)^{-(k+1)/2} \exp\left\{y^2 - \frac{(y-2xt)^2}{1-4t^2}\right\} H_k\left\{\frac{y-2xt}{\sqrt{1-4t^2}}\right\} \end{aligned} \tag{3.24}$$

which has been derived by the Weisner method [7].

3.3. Laguerre polynomials. We define the Laguerre polynomials by means of the series [1]

$$L_n^a(z) = \sum_{m=0}^{n} (-)^m \frac{(1+a)_n}{(1+a)_m} \frac{z^m}{(n-m)!m!}. \tag{3.25}$$

Replacing $z^m/m!$ by $D^{n-m}z^n/n!$ and $(1+a)_n/(1+a)_m$ by $(-)^{n-m}(-n-a)_{n-m}$ in (3.25), one gets:

$$\begin{aligned} L_n^a(z) &= \sum_{m=0}^{n} \frac{(-)^n(-n-a)_{n-m}D^{n-m}}{(n-m)!} \frac{z^n}{n!} \\ &= (-)^n(1-D)^{n+a}z^n/n!. \end{aligned} \tag{3.26}$$

The representation (3.26) is very useful for the derivation of differential recurrence relations of Laguerre polynomials. By similar manipulations, one obtains the following formulae, which are more suitable than (3.26) for the calculation of generating functions:

(3.27) $$z^n L_n^a\left(\frac{1}{z}\right) = (1+a)_n \, {}_0F_1(-; 1+a; -D)z^n/n!$$

i.e.

(3.28) $$L_n^a(z) = (1+a)_n z^n \, {}_0F_1(-; 1+a; z^2D)z^{-n}/n!.$$

More generally, we have

(3.29) $$z^n \, {}_{p+1}F_q(-n, a_1, \cdots, a_p; b_1, \cdots, b_q; z^{-1}) = {}_pF_q(a_1, \cdots, a_p; b_1, \cdots, b_q; -D)z^n.$$

3.4. Gegenbauer polynomials. Let us calculate the polynomials which satisfy the following recurrence relation:

(3.30) $$[(1-z^2)D + nz]y_n = (n+\alpha)y_{n-1}.$$

Using the factorization formula (2.11), one can write instead of (3.30)

(3.31) $$(1-z^2)^{n/2+1}D(1-z^2)^{-n/2}y_n = (n+\alpha)y_{n-1}.$$

Iterating (3.31) n times and putting $y_0 = 1$, one gets

(3.32) $$[(1-z^2)^{3/2}D]^n(1-z^2)^{-n/2}y_n = (1+\alpha)_n.$$

Writing now

(3.33) $$u = z(1-z^2)^{-1/2} \quad \text{and} \quad \hat{D} \equiv d/du = (1-z^2)^{3/2}D,$$

we have after some calculations

(3.34) $$y_n = (1+u^2)^{-n/2}\hat{D}^{-n}(1+\alpha)_n = (1+\alpha)_n(1+u^2)^{-n/2}F(\hat{D})u^n/n!,$$

where, according to (2.3), $F(z)$ is an arbitrary analytic function satisfying $F(0) = 1$.

From (2.13), one can put (3.34) into the form

$$y_n = (1+\alpha)_n F\left(\hat{D} + \frac{nu}{1+u^2}\right)u^n(1+u^2)^{-n/2}/n!$$

i.e.

(3.35) $$y_n = (1+\alpha)_n F(A)z^n/n!$$

where

(3.36) $$A = (1-z^2)^{3/2}D + nz(1-z^2)^{1/2}.$$

But, from (2.11) one can write

$$A^k = (1-z^2)^{n/2}(1-z^2)^{3/2}D^k(1-z^2)^{-n/2}$$

so that

(3.37) $$A^k z^n = (1-z^2)^{n/2}\hat{D}^k u^n = (1-z^2)^{k/2}D^k z^n.$$

Introducing now the set of operators

(3.38) $$B_k = (1-z^2)^{k/2}D^k$$

we have $A^k z^n = B_k z^n$ although B_k does not depend on n. That property allows us to put (3.35) into the form of a symbolic relation [10]

$$y_n \doteqdot (1+\alpha)_n F(B) z^n/n! \tag{3.39}$$

where the undefined symbols B^k in the expansion of $F(B)$ must be replaced by the well defined operators B_k. Equation (3.39) is the most general solution of (3.30). It clearly shows that the expansion coefficients of $F(B)$ are related to the set $y_n(0)$ by the relation

$$F_n = (1+\alpha)_n^{-1} y_n(0) \tag{3.40}$$

so that

$$F(B) = \sum_{n=0}^{\infty} (1+\alpha)_n^{-1} y_n(0) B_n. \tag{3.41}$$

The Jacobi polynomials $P_n^{(\alpha,\alpha)}(z)$ [4] satisfy (3.30). For these polynomials one gets:

$$F(B) \doteqdot {}_2F_1(-;1+\alpha;-B^2/4). \tag{3.42}$$

This leads to the following representation of Gegenbauer polynomials:

$$C_n^\lambda(z) \doteqdot (2\lambda)_n\, {}_0F_1(-;\lambda+1/2;-B^2/4) z^n/n!. \tag{3.43}$$

3.5. Sheffer polynomials. The polynomials of Sheffer A-type zero [10] are a simple set of polynomials having the property

$$J(D) P_n(z) = P_{n-1}(z) \tag{3.44}$$

where $J(z)$ is an entire function satisfying $J(0)=0$, i.e. $J(z) = zj(z)$ with $j(0) \neq 0$.

From (3.44) we obtain

$$J^n(D) P_n(z) = P_0(z) = C \tag{3.45}$$

so that from Eq. (2.11):

$$\begin{aligned} P_n(z) &= J^{-n}(D) C \\ &= j^{-n}(D) D^{-n} C \\ &= F(D) j^{-n}(D) z^n/n!; \qquad F(0) = C. \end{aligned} \tag{3.46}$$

Equation (3.46) is the general representation of polynomials of Sheffer A-type zero. We note that this equation may also be obtained from the Rodrigues formula for polynomials of binomial types derived by Rota et al. [11].

5. Conclusions and acknowledgments. We think that the results obtained in this work are interesting chiefly from the viewpoint of methodology. These results show for instance that in the domain of special functions, long series manipulations may be replaced by the synthetical method of differintegral hyperdifferential operators. We hope that there will be more studies on these operators in order to calculate more expressions $F(z, D) f(z)$. This will surely facilitate the utilizations of special functions.

The author is indebted to Professor M. Demeur, Professor R. Dagonnier and Dr. G. Reidemeister for numerous remarks and discussions.

REFERENCES

[1] M. ABRAMOWITZ AND I. A. SEGUN, *Handbook of Mathematical Functions*, Dover, New York, 1968.

[2] V. BARGMAN, *On the representations of the rotation group*, Rev. Mod. Phys., 34 (1962), pp. 829–845.

[3] R. COURANT AND D. HILBERT, *Methods of Mathematical Physics*, vol. II, Interscience, New York, 1962.

[4] A. ERDÉLYI, W. MAGNUS, F. OBERHETTINGER, AND F. G. TRICOMI, *Tables of higher Transcendental Functions and Tables of Integral Transforms*, McGraw-Hill, New York, 1954.

[5] H. HOLMGREN, *Sur l'intégration de l'équation différentielle* $(a_2+b_2x+C_2x^2)\,d^2y/dx^2+(a_1+b_1x)\,dy/dx+a_0y=0$, Kongliga Svenska Ventenkaps—Academiens, 7 (1867–1868), no. 9.

[6] J. LIOUVILLE, *Mémoire sur l'intégration de l'équation* $(mx^2+nx+p)\,d^2y/dx^2+(qx+r)\,dy/dx+sy=0$ *à l'aide des différentielles à indices quelconques*, J. Ecole Polytech., 13 (1832), pp. 163–186.

[7] E. B. MCBRIDE, *Obtaining Generating Functions*, Springer-Verlag, New York, 1971.

[8] M. MILLER AND S. STEINBERG, *Applications of Hyperdifferential Operations to Quantum Mechanics*, Comm. Math. Phys., 24 (1971), 40–60.

[9] K. B. OLDHAM AND J. SPANIER, *The Fractional Calculus*, Academic Press, New York and London, 1974.

[10] E. D. RAINVILLE, *Special Functions*, Macmillan, New York, 1965.

[11] G. C. ROTA, D. KAHANER AND A. ODLYZKO, *On the foundations of combinatorial theory VIII. Finite operator calculus*, J. Math. Anal. Appl. 42 (1973), pp. 684–760.

[12] F. TREVES, *Hyperdifferential operators in complex space*, Bull. Soc. Math. France, 97 (1969), pp. 193–223.

[13] R. M. WILCOX, *Exponential operators and parameter differentiation in quantum physics*, J. Mathematical Phys., 8 (1967), pp. 962–982.

[14] K. B. WOLF, *Canonical transforms*, Ibid., 15 (1974), pp. 1295–1301.

[15] ———, *Hyperdifferential operators and integral transforms*, Rev. Mexicana Fis., (1976), March.

SIAM J. MATH. ANAL.
Vol. 9, No. 6, December 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0007 $01.00/0

A METHOD OF GLOBAL BLOCKDIAGONALIZATION FOR MATRIX-VALUED FUNCTIONS*

H. GINGOLD†

Abstract. Let $A(x)$ be an $n \times n$ analytic matrix function of the vector variable x. Let the eigenvalues of $A(x)$ belong to two disjoint sets for every fixed x. Then there exists an invertible analytic matrix function $M(x)$ which takes $A(x)$ by a similarity transformation into a blockdiagonal form. Similar theorems for $A(x)$ being smooth are also proved.

1. Introduction. Every constant $n \times n$ matrix with entries in the field of complex numbers and having more than one distinct eigenvalue is similar, in that field, to a blockdiagonal matrix of two diagonal blocks that have no common eigenvalues. If that matrix, $A = A(x)$, is a function of x, where x is in a domain D of $\mathbb{R}^m$ or $\mathbb{C}^m$, this is still true for every $x \in D$, provided the eigenvalues can be labeled so that the numbers $\lambda_1(x), \cdots, \lambda_k(x)$ are distinct from $\lambda_{k+1}(x), \cdots, \lambda_n(x)$ for all $x \in D$. Thus, there exists an identity of the form

$$M^{-1}(x)A(x)M(x) = A_1(x) \oplus A_2(x), \qquad x \in D, \tag{1.1}$$

where $A_1(x)$, $A_2(x)$ are $k \times k$ and $(n-k)\times(n-k)$ matrices, respectively.

The purpose of this investigation is to replace the trivial result (1.1) by a more precise one which tells how "smooth" $M(x)$ can be taken when $A(x)$ has some known smoothness properties, such as being in $C^m(D)$ or holomorphic.

A method of solution of the above problem is provided by Sibuya [10] and Hsieh and Sibuya [6]. A closely related problem to the above is solved in Wasow [12].

Our approach to characterize $M(x)$ hinges on the idea of making $M(x)$ a solution of an ordinary linear homogeneous matrix differential equation. Thanks to that, the existence and "smoothness" of $M(x)$ and $M^{-1}(x)$ is taken care of simultaneously in the *whole* domain of interest.

To achieve this we link our problem to the special case of projection matrices. See Coppel [2] for a similar approach in the case that x is one-dimensional variable.

The possibility of finding a smooth $M(x)$ in (1.1) plays a vital role in the process of simplification of linear singular differential systems. For example, see Wasow [13, p. 138].

2. Preliminaries.

Notation 2.1. We denote by x an r-dimensional vector variable $x = (x_1, \cdots, x_r)$, $1 \leqq r$ in a set D to be defined. We denote by $|x|$ a norm on x.

Notation 2.2. We have $D = [a, b]^r$ or $D = K^r$, where K is a finite simply connected domain in the complex plane.

Notation 2.3. By $A(x) \in C^m(D)$, $m \geqq 0$ we mean $A(x) \in C^m([a, b]^r)$ or $A(x)$ is analytic in K^r.

Notation 2.4. We set $P_k = \operatorname{diag}(1, \cdots, 1, 0, \cdots, 0)$, k times 1 and $n-k$ times 0.

Notation 2.5. "Blockdiagonal" shall mean a decomposition such as in (1.1) *with fixed* k, independent of x.

For the sake of simplicity we will sometimes suppress the variable x in the matrix notation.

* Received by the editors September 9, 1976, and in final revised form May 13, 1977.
† Department of Mathematics, University of Utah, Salt Lake City, Utah. 84112.

We list some facts which will be needed in the sequel. Since they are so simple, their proofs will not be elaborated on here.

(I) *M is blockdiagonal iff $MP_k = P_kM$.*

(II) *If M and N are blockdiagonal and N is invertible, then $N^{-1}MN$ is blockdiagonal.*

(III) *If $P = P(x)$ is differentiable in D, $D \in \mathbb{R}$ or $D \in \mathbb{C}$, and $P^2 = P$, then*

$$(P^2)' = P' = P'P + PP', \qquad \left(' = \frac{d}{dx}\right),$$

(IV) $PP'P = 0$.

(V) *If M is invertible then MP_kM^{-1} is a projection.*

(VI) *Every projection matrix is similar to a diagonal projection matrix.*

(VII) *$M^{-1}AM$ is blockdiagonal iff A commutes with $P := MP_kM^{-1}$.*

Property (VII) reduces the blockdiagonalization problem for $A(x)$ to one for projection matrices.

Hypothesis H_1. Let $\Phi(x, \lambda) = \det(A - \lambda I)$. Then $\Phi(x, \lambda) = \Phi_1(x, \lambda)\Phi_2(x, \lambda)$, where Φ_1 and Φ_2 are relatively prime monic polynomials in λ of *constant* degrees k and $n-k$, respectively, for *all* $x \in D$, with coefficients having the same regularity properties as $A(x)$ (namely belonging to $C^m(D)$).

Notice that Hypothesis H_1 means that the eigenvalues of $A(x)$ may be labeled so that the eigenvalues $\lambda_1(x), \cdots, \lambda_k(x)$ are distinct from the eigenvalues $\lambda_{k+1}(x), \cdots, \lambda_n(x)$.

From now on we identify the spectra of Φ_1 and Φ_2 with the eigenvalues $\lambda_1(x), \cdots, \lambda_k(x)$ and $\lambda_{k+1}(x), \cdots, \lambda_n(x)$ correspondingly.

3. A lemma.

LEMMA 3.1. *Let $A(x) \in C^m(D)$ and let Hypothesis H_1 hold. Then for every x there exist in the λ-plane a finite set of rectifiable closed Jordan curves which will be denoted by Γ_x. This set Γ_x, contains the spectra of Φ_1 in its interior and has the spectra of Φ_2 in its exterior.*

The matrix function $P(x)$ defined by

$$P(x) = \frac{1}{2\pi i} \int_{\Gamma_x} [\lambda I - A(x)]^{-1} \, d\lambda \tag{3.1}$$

belongs to $C^m(D)$. Moreover $P(x)$ is a projection which commutes with $A(x)$.

Proof. Observe that;

i) If $A(x)$ is a continuous matrix function of x, $\lambda_i(x)$ is also continuous in x for every i, $i = 1, 2, \cdots, n$. Therefore:

ii)

$$\inf_{\substack{x \in D \\ 1 \leq i \leq k \\ k < j \leq n}} |\lambda_i(x) - \lambda_j(x)| = \delta > 0 \tag{3.2}$$

and there exists h_0, such that

$$|\Delta x| \equiv |x - \hat{x}| < h_0 \quad \text{implies} \quad |\lambda_i(x) - \lambda_i(\hat{x})| < \delta/5 \quad \text{for } i = 1, 2, \cdots, n. \tag{3.3}$$

(We denote $\Delta x \equiv (\Delta x_1, \cdots, \Delta x_n)$.) Also denote

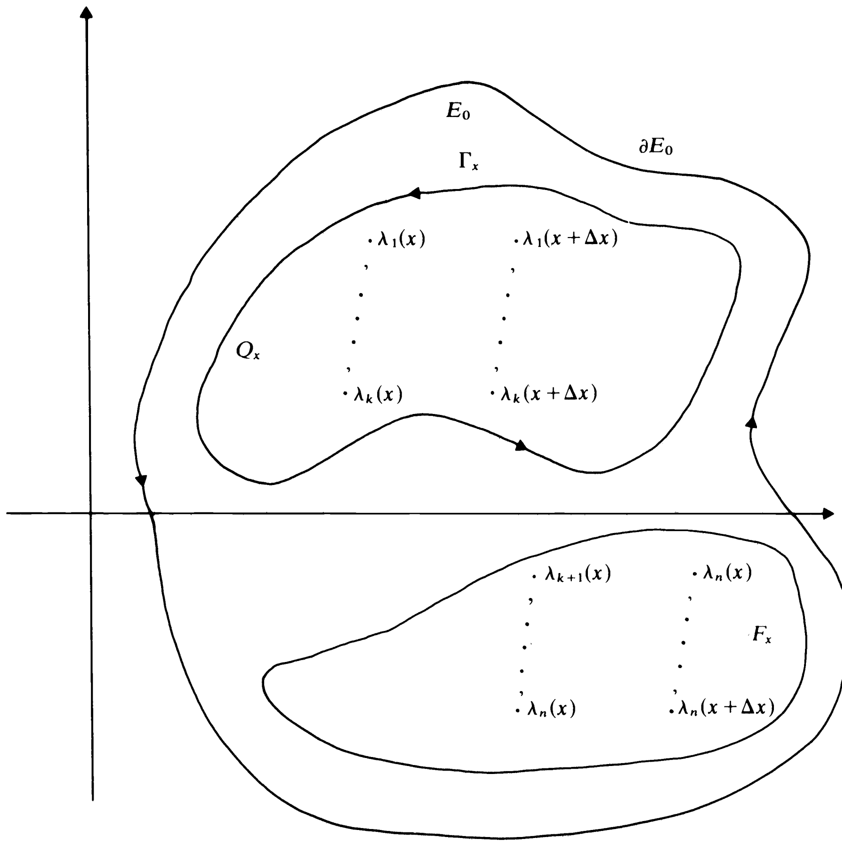

FIG. 1. *The two disjoint sets of eigenvalues in the* λ*-plane at the "time"* x *and at the "time"* $x+\Delta x$.

$$E_0=\{\lambda\mid |\lambda|<\max_{i,x}|\lambda_i(x)|+2\delta,\ i=1,\cdots,n,\ x\in D\},$$

$$F_x=\bigcup_{j=k+1}^{j=n}F_{jx},\qquad F_{jx}=\{\lambda\mid |\lambda-\lambda_j(x)|\leqq 2\delta/5\},$$

$$Q_x=\bigcup_{i=1}^{i=k}Q_{ix},\qquad Q_{ix}=\{\lambda\mid |\lambda-\lambda_i(x)|\leqq 2\delta/5\}.$$

Obviously: $F_x\cap Q_x=\varnothing$, $(E_0\backslash F_x)$ is open, Q_x is closed, and $Q_x\subset E_0\backslash F_x$. (See Fig. 1.)

By [11] we are guaranteed that if Q_x is closed and $(E_0\backslash F_x)$ is open and bounded in the complex λ-plane, $Q_x\subset(E_0\backslash F_x)$, then there exists a Cauchy domain R_x, such that $Q_x\subset R_x$ and $\bar{R}_x\subset(E_0\backslash F_x)$, ($\bar{R}_x$ is the closure of R_x), and therefore there exists a finite positive number of rectifiable Jordan closed curves which will have in their interior the domain Q_x, and since $F_x\cap Q_x=\varnothing$, $F_x\cap(E_0\backslash F_x)=\varnothing$, $\lambda_{k+1}(x),\cdots,\lambda_n(x)$ will be outside of them. For any $x\in D$, choose Γ_x in (3.1) to be identical with ∂R.

Recall: R is a Cauchy domain if:

i) R is bounded and open;

ii) R has a finite number of components the intersection of two of which is disjoint;

iii) the boundary ∂R, is composed of a finite number of closed rectifiable Jordan curves, no two of which intersect.

If we consider the same domains in the λ-plane for the point $x+\Delta x$, with $|\Delta x|<h_0$, and examine the location of $\lambda_1(x+\Delta x), \cdots, \lambda_k(x+\Delta x), \lambda_{k+1}(x+\Delta x), \cdots, \lambda_n(x+\Delta x)$, then $|\Delta x|<h_0$, where h_0 was chosen in (3.3), implies that every point $\lambda_i(x)$, $i=1,2,\cdots,n$, has a distance greater than $\delta/5$ from $\Gamma_{x+\Delta x}$, and every point $\lambda_i(x+\Delta_x)$, $i=1,2,\cdots,n$, has a distance from Γ_x greater than $\delta/5$. Therefore

$$\begin{aligned} P(x)&=\frac{1}{2\pi i}\oint_{\Gamma_x}[\lambda I-A(x)]^{-1}\,d\lambda \\ &=\frac{1}{2\pi i}\oint_{\Gamma_{x+\Delta x}}[\lambda I-A(x)]^{-1}\,d\lambda=\frac{1}{2\pi i}\oint_{\Gamma}[\lambda I-A(x)]^{-1}\,d\lambda \end{aligned} \tag{3.4}$$

for every Γ which is a set of rectifiable Jordan closed curves encircling in its interior, $\lambda_1(x),\cdots,\lambda_k(x)$, and in its exterior, $\lambda_{k+1}(x),\cdots,\lambda_n(x)$. This is true by Cauchy's integral formula. In the forthcoming formulas (3.5), (3.6), (3.7) when computing partial derivatives we will assume $\Delta x=(0,\cdots,0,\Delta x_i,0,\cdots,0)$.

We have for every $i=1,\cdots,r$,

(3.5)
$$\frac{P(x+\Delta x)-P(x)}{\Delta x_i}=\frac{1}{2\pi i}\left[\oint_{\Gamma_{x+\Delta x}}[\lambda I-A(x+\Delta x)]^{-1}\,d\lambda-\oint_{\Gamma_x}[\lambda I-A(x)]^{-1}\,d\lambda\right]\frac{1}{\Delta x_i}$$

$$=\frac{1}{2\pi i}\oint_{\Gamma_x}\frac{\{[\lambda I-A(x+\Delta x)]^{-1}-[\lambda I-A(x)]^{-1}\}\,d\lambda}{\Delta x_i}. \tag{3.6}$$

This implies

$$P'(x)=\frac{\partial}{\partial x_i}P(x)=\frac{1}{2\pi i}\oint_{\Gamma_x}d\lambda\frac{\partial}{\partial x_i}[\lambda I-A(x)]^{-1}. \tag{3.7}$$

Moreover, for $|\Delta x|<h_0$,

$$\begin{aligned} P'(x+\Delta x)-P'(x)&=\frac{1}{2\pi i}\oint_{\Gamma_{x+\Delta x}}d\lambda\frac{\partial}{\partial(x+\Delta x)}[\lambda I-A(x+\Delta x)]^{-1} \\ &\quad-\frac{1}{2\pi i}\oint_{\Gamma_x}d\lambda\frac{\partial}{\partial x_i}[\lambda I-A(x)]^{-1} \\ &=\frac{1}{2\pi i}\oint_{\Gamma_x}d\lambda\frac{\partial}{\partial(x+\Delta x)}[\lambda I-A(x+\Delta x)]^{-1}-\frac{\partial}{\partial x}[\lambda I-A(x)]^{-1}. \end{aligned} \tag{3.8}$$

This proves that if $A(x)\in C^1(D)$ then $P(x)\in C^1(D)$. By applying an induction argument we obtain $A(x)\in C^m(D)$ implies $P(x)\in C^m(D)$. The fact that $P(x)$ is a projection follows from [8, p. 419].

From now on in this paper $P(x)$ will stand for the projection defined in Lemma 3.1 and given by (3.1).

4. The one-dimensional case. We assume throughout this section that our domain D is one-dimensional or $r=1$.

LEMMA 4.1. *Let $P=P(x)$ be a projection matrix, and $P\in C^m(D)$, $m>0$. If $P(x)$ is similar to P_k at one point $x_0\in D$, it is similar to P_k in all of D, and among the matrices $W(x)$ such that $W^{-1}(x)P(x)W(x)=P_k$, $x\in D$, there is one which is a solution of the differential equation*

$$W'=(P'P-PP')W, \tag{4.1}$$

with $W(x)\in C^m(D)$.

Proof. For any invertible solution of (4.1) one has

$$W'W^{-1}=P'P-PP' \tag{4.2}$$

and therefore, using properties (III) and (IV),

$$(W'W^{-1})P-P(W'W^{-1})=P'; \tag{4.3}$$

by left and right multiplication with W^{-1} and W, respectively, this becomes

$$-W^{-1}W'W^{-1}PW+W^{-1}P'W+W^{-1}PW'=0, \tag{4.4}$$

i.e.

$$\frac{d}{dx}(W^{-1}PW)=0. \tag{4.5}$$

By assumption, there is a constant matrix W_0 such that $W_0^{-1}P(x_0)W_0=P_k$. Let $W=W(x)$ be the unique fundamental solution of (4.1) characterized by the initial condition $W(x_0)=W_0$. Then $W^{-1}(x_0)P(x_0)W(x_0)=P_k$ and since $W^{-1}(x)P(x)W(x)$ is constant, by (4.5)

$$W_0^{-1}(x)P(x)W(x)=P_k, \quad \text{for all } x\in D. \tag{4.6}$$

As a solution of a linear differential equation with coefficients in $C^{m-1}(D)$, the matrix function $W(x)$ is in $C^m(D)$, as was to be proved. The above lemma leads us to the next theorem.

THEOREM 4.1. *Let $A(x)\in C^m(D)$ and Hypothesis H_1 be satisfied. Then*:

i) *There exists an invertible matrix function $M(x)$, $M(x)\in C^m(D)$ such that*

$$M^{-1}(x)A(x)M(x)=A_1(x)\oplus A_2(x). \tag{4.7}$$

The eigenvalues of $A_1(x)$ and $A_2(x)$ are the roots of $\Phi_1(x,\lambda)$ and $\Phi_2(x,\lambda)$ correspondingly.

ii) *Moreover, if $A(x)$ is periodic with period τ,* (4.7) *may be satisfied with a periodic $M(x)$ with period τ.*

Proof. By Lemma 4.1 we find an invertible $W(x)$, $W(x)\in C^m(D)$ such that (4.6) holds.

i) We choose $M(x)=W(x)$. Use property (I) to find out that (4.7) is satisfied. Moreover, to the projection P defined by (3.1) there corresponds a unique decomposition of the vector space on which $A(x)$ operates into a direct sum of two disjoint subspaces. Therefore, the eigenvalues of A_1 and A_2 are the roots of Φ_1 and Φ_2 correspondingly.

ii) By (3.1), $P'(x)$ and $P(x)$ are also periodic with period τ whenever $A(x)$ is such. In addition to (4.6) we also have

$$W^{-1}(x+\tau)P(x)W(x+\tau)=P_k. \tag{4.8}$$

Eliminate $P(x)$ from (4.6) and (4.8) to obtain

$$W^{-1}(x)W(x+\tau)P_k=P_kW^{-1}(x)W(x+\tau). \tag{4.9}$$

This means that $W^{-1}(x)W(x+\tau)$ is blockdiagonal.

Since $W(x)$, $W(x+\tau)$ are solutions of the same differential equation (4.1),

$$W(x+\tau)=W(x)C \tag{4.10}$$

where C is a constant invertible matrix and C was shown in (4.9) to be blockdiagonal.

We are looking for a matrix $N(x)$ such that

$$W(x+\tau)N(x+\tau)=W(x)N(x). \tag{4.11}$$

By (4.10) it follows that

$$CN(x+\tau)=N(x). \tag{4.12}$$

Many matrices satisfy (4.12) and consequently (4.11). For example,

$$N(x)=\exp\left[-\frac{x}{\tau}\log C\right]. \tag{4.13}$$

Now choose $N(x)$ as in (4.13) and $M(x)=W(x)N(x)$ to obtain the desired transformation.

Remark. The differential equation (4.1) appears in Kato's paper [7] with "no justification." We can easily show how (4.1) may be derived in a natural way by the blockdiagonalization problem. By property (VII) our problem was reduced to finding an invertible matrix $W\in C^m(D)$ such that $W^{-1}PW=P_k$. We go in *inverse* order through the steps (4.3), (4.4), (4.5) to find out that $W'W^{-1}$ must satisfy (4.3). In order to pass from (4.3) to (4.2) we can easily verify that if E, F, L are $n\times n$ matrices such that $E^2=E$, $F^2=F$, $LE+FL=L$, then a particular solution to the matrix equation $YE-FY=L$ is given by $Y=LE-FL$. See also Rosenblum [9].

5. The *r*-dimensional case.

THEOREM 5.1. *Assume $A(x)\in C^m(D)$, $m\geqq 2$, and let Hypothesis* H_1 *be satisfied. Then there exists an invertible matrix $M(x)\in C^{m-1}(D)$ such that*

$$M^{-1}(x)A(x)M(x)=A_1(x)\oplus A_2(x).$$

The eigenvalues of $A_1(x)$ and $A_2(x)$ are the roots of $\Phi_1(x,\lambda)$ and $\Phi_2(x,\lambda)$ correspondingly.

Proof. We first adopt a couple of notations.

Notation 5.1. i) Define D_ν to be the restriction of the set D to the variables $x_1,\cdots,x_\nu$, $\nu=1,2,\cdots,r$. $D_\nu=\{x_1,\cdots,x_\nu)|x\in D\}$.

ii) Denote

$$\begin{aligned}
&P_0=P(x^0),\\
&P_\nu=P(x_1,x_2,\cdots,x_\nu,x^0_{\nu+1},\cdots,x^0_r), \qquad \nu=1,\cdots,r,\\
&P_r=P(x),\\
&P'_\nu=\frac{\partial}{\partial x_\nu}P_\nu.
\end{aligned} \tag{5.1}$$

iii) Define the sequence W_ν, $\nu=0,1,\cdots,r$, as follows: W_0 is the matrix satisfying

$$W_0^{-1}P_0W_0=P_k; \tag{5.2}$$

and W_ν is the unique solution of the equation

$$W'_\nu=(P'_\nu P_\nu-P_\nu P'_\nu)W_\nu, \qquad \frac{\partial W_\nu}{\partial x_\nu}=W'_\nu \tag{5.3}$$

satisfying

$$W_\nu=W_\nu(x_1,x_2,\cdots,x_{\nu-1},x_\nu,x^0_\nu,W_{\nu-1}) \tag{5.4}$$

where: $x_1, \cdots, x_{\nu-1}$, are considered as parameters, x_ν is the independent variable, and W_ν takes the "initial value" $W_{\nu-1}$ at the point x_ν^0, and $W_{\nu-1}$ itself depends on the parameters $x_1, \cdots, x_{\nu-1}$.

We proceed now by induction to show that: There exist W_ν which are invertible and such that

$$W_\nu^{-1} P_\nu W_\nu \equiv P_k, \qquad \nu = 1, 2, \cdots, r. \tag{5.5}$$

For $\nu = 0$ it is true by taking P_0 into its Jordan canonical form. Assume $W_{\nu-1}$ to satisfy induction hypothesis, and proceed to W_ν. For any fixed point $(x_1, \cdots, x_{\nu-1}) \in D_{\nu-1}$, W_ν must be an invertible matrix function for all x_ν since we chose it to be invertible at x_0. This is true by [1, p. 28].

For the smoothness properties of W_ν we notice that $W_\nu \in C^m(\hat{D})$ for every fixed $(x_1, \cdots, x_{\nu-1}) \in D_{\nu-1}$, where $\hat{D}$ is the domain of variation of x_ν. By [5, Chap. 3], $W_\nu \in C^{m-1}(D_{\nu-1})$ since the coefficient matrix in (5.3) and the "initial value" belong to $C^{m-1}(D_{\nu-1})$. Therefore $W_\nu \in C^{m-1}(D_\nu)$ for $\nu = 1, \cdots, r$. We choose $M(x) = W_r(x)$ and the result follows.

Remark. It is impossible to extend part ii) of Theorem 4.1 to the case $r > 1$. This was pointed out in [6] by Hsieh and Sibuya.

Remark. The properties of continuity and differentiability inherited from $P'_\nu P_\nu - P_\nu P'_\nu$ by $W_\nu(x)$ can be easily deduced by writing down a fundamental solution of (5.3) in terms of the resolvent series.

Remark. In the case that x is a scalar we obtained $M(x) \in C^m(D)$. In the case $r > 1$ we obtained $M(x) \in C^{m-1}(D)$. The difference stems from the fact that the solution of a differential equation depending on parameters in the coefficients is usually not smoother with respect to the parameters than are the given coefficients. The theorems in [6] by Hsieh and Sibuya are stronger in the sense that $A(x) \in C^m(D)$ implies the existence of $M(x) \in C^m(D)$ for $m \geqq 0$.

Acknowledgment. I hereby wish to thank Professor W. A. Harris Jr. for his encouragement and enthusiasm.

REFERENCES

[1] E. A. CODDINGTON AND N. LEVINSON, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[2] W. A. COPPEL, *Dichotomies and reducibility*, J. Differential Equations, 3 (1967), pp. 500–521.

[3] N. DUNFORD AND J. T. SCHWARTZ, *Linear Operators Part I: General Theory*, Interscience, New York, 1958.

[4] F. R. GANTMACHER, *The theory of Matrices*, Chelsea, New York, 1966.

[5] E. HILLE, *Lectures on Differential Equations*, Addison-Wesley, Reading, MA, 1969.

[6] P. F. HSIEH AND Y. SIBUYA, *A global analysis of matrices of functions of several variables*, J. Math. Anal. Appl., 14 (1966), pp. 332–340.

[7] T. KATO, *Notes on perturbation theory*, Bull. Amer. Math. Soc., 61 (1955), p. 146.

[8] F. RIESZ AND B. SZ-NAGY, *Functional Analysis*, Frederick Ungar, New York, 1955.

[9] M. ROSENBLUM, *On the operator equation $BX - XA = Q$*, Duke Math. J., 23 (1956), pp. 263–270.

[10] Y. SIBUYA, *Some global properties of matrices of functions of one variable*, Math. Ann., 161 (1965), pp. 67–77.

[11] A. E. TAYLOR, *Spectral theory of closed distributive operators*, Acta Math., 84 (1951), pp. 189–224.

[12] W. WASOW, *On holomorphically similar matrices*, J. Math. Anal. Appl., 4 (1962), pp. 202–206.

[13] ———, *Asymptotic Expansions for Ordinary Differential Equations*, Wiley-Interscience, New York, 1965.

SIAM J. MATH. ANAL.
Vol. 9, No. 6, December 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0008 $01.00/0

THE GENERALIZED INVERSE OF AN UNBOUNDED LINEAR OPERATOR WITH UNBOUNDED CONSTRAINTS*

W. F. LANGFORD†

Abstract. The concept of best least squares solution is used to define a generalized inverse of an unbounded linear operator between two inner product spaces, subject to unbounded linear functional constraints. The nullspace of the operator is assumed finite dimensional. A necessary and sufficient condition for the existence of this generalized inverse is established. When the condition holds, the calculation of the best least squares solution is reduced to an explicit algebraic formula. The theory is illustrated by application to a general linear two-point boundary value problem, for which a new proof of the existence and uniqueness of the best least squares solution is obtained, without the use of Green's functions.

1. Introduction. Motivated by applications to boundary value problems, let us consider the following situation. A linear mapping L takes an inner product space X onto an inner product space Y. The spaces X and Y need not be complete, and L need not be bounded, but L is assumed to have a finite dimensional nullspace $N(L)$. The algebraic duals of X and Y are denoted X^* and Y^*; these are spaces of all linear functionals, not necessarily bounded, on X and Y respectively. A subset $\{f_1, \cdots, f_m\}$ of X^* and an element y of Y together define an interesting problem: among all x in X satisfying the *constraints*

$$f_i(x)=0, \qquad i=1, \cdots, m, \tag{1}$$

find the best least squares of the equation

$$Lx=y. \tag{2}$$

A precise definition of best least squares solution is given in the next section. Henceforth it is abbreviated to BLSS and denoted by $x^\dagger$.

It is not surprising that in some cases problem (1) (2) has no exact solution. In fact, it can happen that problem (1) (2) does not have even a BLSS. Let us define the functionals $f_1, \cdots, f_m$ to be *admissible constraints for* L if and only if problem (1) (2) has a BLSS for any y in Y. In §§ 3 and 5 we give necessary and sufficient conditions for a set of functionals to be admissible constraints for L.

Let $F\equiv[f_1, \cdots, f_m]$ be the subspace of X^* spanned by $f_1, \cdots, f_m$, and let $F_0=\{x\in X \mid f(x)=0 \text{ for all } f\in F\}$ be the annihilator of F in X. Then F_0 is a subspace of X consisting precisely of those x satisfying (1). Now define $\mathscr{L}=L|_{F_0}$, the restriction of L to F_0. In case the BLSS $x^\dagger$ exists and is unique for each $y\in Y$, there is defined a unique mapping from Y to F_0, called the *generalized inverse of* $\mathscr{L}$ and denoted $\mathscr{L}^\dagger$, that is $\mathscr{L}^\dagger y=x^\dagger$. Thus, finding the BLSS of problem (1) (2) for every $y\in Y$ is equivalent to calculating the generalized inverse $\mathscr{L}^\dagger$.

Many other generalized inverses can be defined [2], but it is claimed that the type of inverse considered here has great practical importance. This $\mathscr{L}^\dagger$ is the analogue of the Moore–Penrose generalized inverse of linear algebra. Besides the obvious relevance to least squares approximation problems, our generalized inverse arises naturally in bifurcation theory. Very briefly: bifurcation is the branching of solutions to a problem. The classical inverse function theorem guarantees local uniqueness.

* Received by the editors October 23, 1975, and in revised form September 20, 1976.
† Department of Mathematics, McGill University, Montreal, Canada H3A 2K6.

Therefore bifurcation can occur only when the inverse function theorem fails to apply, and so bifurcation theory can be thought of as a study of "generalized inverse function theorems." The generalized inverse studied here is most appropriate for many bifurcation problems; see [6], [7], [12].

The next section contains basic definitions necessary for the study of problem (1) (2). Section 3 presents the main result of this paper, an existence and uniqueness theorem for the BLSS of problem (1) (2). In § 4, this theory is applied to the general linear two-point boundary value problem

$$\frac{dx}{dt}+Ax=y(t), \qquad t\in(a,b), \tag{3}$$

$$Mx(a)+Nx(b)=0, \tag{4}$$

where $x\in X=C_n^1[a,b]$, $y\in Y=C_n[a,b]$, and M and N are $m\times n$ matrices with $0\leqq m\leqq 2n$. Previous studies [1], [8], [11] of this problem used a cumbersome process of completion to the Hilbert space $L^2[a,b]$, eventually followed by the proof of regularity results using generalized Green's functions. Here we are able to show directly that there exists a unique BLSS $x^\dagger$ in $C_n^1[a,b]$.

Section 5 presents useful explicit formulae which characterize the set of admissible constraints, and facilitate the calculation of the BLSS. Finally, the Appendix gathers together certain identities which are used in proving the main theorems.

2. Definitions. All scalars are assumed real for convenience; the extension to complex inner product spaces is straightforward. The assumption that L maps onto Y is not a limitation in applications. Since Y need not be complete, it can be defined to be the range $R(L)$ of the unconstrained mapping L.

The two inner products in X and Y are both denoted by the same symbol (u,v) without confusion, and the two resulting norms are indicated by the usual notation $\|u\|$. The operation of a functional f in X^* on x in X is represented by the notations $\langle x,f\rangle=f(x)$, and similarly for Y^* and Y. The spaces of *bounded* linear functionals on X and Y (in the topology induced by the inner product norms) are denoted X' and Y' respectively, and called here the *normed duals* of X and Y. Most recent works on best least squares solutions and generalized inverses employ only the normed dual spaces, or in particular Hilbert space (which is its own normed dual), see [2], [9], [10]. Although algebraic duals are very simple in structure, they have important properties which differ subtly from the familiar properties of normed duals. These properties are reviewed in the Appendix.

First we give a precise definition of BLSS (best least squares solution). The set of least squares solutions (LSS) of problem (1) (2) for a given y in Y is

$$S_y=\{u\in F_0\,|\,\|Lu-y\|\leqq\|Lx-y\|\ \forall x\in F_0\}. \tag{5}$$

Then a BLSS $x^\dagger$ of (1) (2) (if it exists) is a minimum norm element of S_y, that is

$$\|x^\dagger\|\leqq\|u\| \quad \forall x\in S_y. \tag{6}$$

The adjoint L^* of L is the mapping of Y^* into X^* defined in the usual way by

$$\langle x, L^*y^*\rangle=\langle Lx, y^*\rangle \quad \forall x\in X, \quad y^*\in Y^*. \tag{7}$$

The adjoint $\mathscr{L}^*$ of $\mathscr{L}$ is defined similarly by

$$\langle x, \mathscr{L}^*y^*\rangle=\langle \mathscr{L}x, y^*\rangle \quad \forall x\in F_0, \quad y^*\in Y^*. \tag{8}$$

The inner product spaces X and Y can be embedded in X^* and Y^* respectively in a natural way, that is given $z \in X$, there is a unique $z^* \in X^*$ defined by

$$\langle x, z^* \rangle = (x, z) \quad \forall x \in X, \tag{9}$$

and similarly for Y and Y^*. It is often convenient to identify such z and z^* and write

$$\langle x, z \rangle = (x, z), \quad x, z \in X. \tag{10}$$

Note that this identification mapping is one-to-one. The notation $S|^{X^*}$ is used to indicate that a subset $S \subset X$ has been embedded in X^*. Sometimes we do not distinguish between a subset of X and its embedded image in X^*, and write for example $X \subset X^*$.

The orthogonal complement of a subset S of an inner product space X is defined by

$$S^\perp = \{x \in X | (s, x) = 0 \, \forall s \in S\}. \tag{11}$$

If S and F are subsets of X and X^* respectively, then the *annihilators* of S and F are defined to be

$$\begin{aligned} S^0 &= \{x^* \in X^* | \langle s, x^* \rangle = 0 \, \forall s \in S\}, \\ F_0 &= \{x \in X | \langle x, f \rangle = 0 \, \forall f \in F\}. \end{aligned} \tag{12}$$

Note that S^0 and F_0 are subspaces but are not closed in general. Annihilators are related to complements by

$$S^\perp = (S|^{X^*})_0, \qquad S^\perp|^{X^*} = S^0 \cap X.$$

3. Existence theory.

LEMMA 1. *Problem* (1) (2) *has a unique BLSS for each* $y \in Y$ *if and only if the annihilator of the range of* $\mathscr{L}$ *lies in* Y, *i.e.*

$$(LF_0)^0 = (LF_0)^\perp \subseteq Y. \tag{13}$$

Proof. For convenience write $Z = (LF_0)^0$. Note Z is finite dimensional since F is and L is onto. First assume (13) is true. Then Z is orthogonally complemented in Y by the corollary to the projection theorem in the Appendix, i.e. $Y = Z \oplus Z^\perp$. But $Z^\perp = [(LF_0)^\perp]^\perp = [(LF_0)^0]_0 = LF_0$, by Lemma 5 in the Appendix. Hence

$$Y = Z \oplus R(\mathscr{L}). \tag{14}$$

Now, by the projection theorem, every $y \in Y$ has a best approximation in $R(\mathscr{L})$, so S_y is nonempty. Finally, S_y has a unique minimum norm element $x^\dagger$ since $N(\mathscr{L})$ is finite dimensional, using another application of the projection theorem.

For the converse, suppose there exists a unique BLSS for every $y \in Y$. Then, by the projection theorem again, for any $y \in Y$ there exist elements $y_1 \in LF_0$ and $y_2 \in (LF_0)^\perp$ such that $y = y_1 + y_2$; therefore

$$Y = LF_0 \oplus (LF_0)^\perp. \tag{15}$$

Suppose (13) is false, then there exists $z_1 \in Z$ with $z_1 \notin (LF_0)^\perp$. Since $(LF_0)^\perp$ is finite dimensional, it has an orthonormal basis $\{u_1, \cdots, u_p\}$. Define $z_2 \in Z$ by

$$\langle y, z_2 \rangle = \langle y, z_1 \rangle - \sum_{k=1}^{p} \langle v_k, z_1 \rangle (y, v_k) \tag{16}$$

for all $y \in Y$. Then

$$\langle v_k, z_2\rangle = 0, \qquad k = 1, \cdots, p, \tag{17}$$

but also $z_2(LF_0) = 0$, so by (15), $z_2 \equiv 0$. But then (16) implies

$$z_1 \in (LF_0)^{\perp}, \tag{18}$$

which completes the proof.

LEMMA 2. *For any F and L as defined in the introduction*

$$L^*(LF_0)^0 = F \cap N(L)^0. \tag{19}$$

Proof. Note $L^*(LF_0)^0 = L^*Z = \{x^* \in X^* | x^*(x) = zLx, z \in Z\}$. If $x \in N(L)$ then $zLx = 0$ for all $z \in Z$, so $L^*Z \subseteq N(L)^0$. If $x \in F_0$, then $zLx = 0$ by definition of Z, so $L^*Z \subseteq (F_0)^0 = F$, using Lemma 6 in the Appendix. Hence $L^*Z \subseteq F \cap N(L)^0$. For the converse, we use $N(L)^0 = R(L^*) = L^*Y^*$ from Lemma 8 in the Appendix. Suppose $x^* \in F \cap N(L)^0$, then $x^* = L^*y^*$ for some $y^* \in Y^*$, and furthermore for all $x \in F_0$, $0 = x^*(x) = y^*Lx$. Hence $y^* \in Z$ and $x^* \in L^*Z$.

THEOREM 1. *Problem* (1) (2) *has a unique BLSS for every $y \in Y$ iff each constraint in F which annihilates $N(L)$ is of the form $f(x) = (Lx, y)$ for some $y \in Y$, i.e.*

$$F \cap N(L)^0 \subseteq L^*Y. \tag{20}$$

Proof. This follows immediately from Lemmas 1 and 2, using the fact that L^* is one-to-one (since L is onto).

A weakened form of Theorem 1, which may be easier to use in applications, follows.

COROLLARY 1. *A sufficient condition for problem* (1) (2) *to have a unique BLSS for each $y \in Y$ is that the constraints* (1) *satisfy*

$$f_i \in N(L)|^{X^*} \oplus L^*Y, \qquad i = 1, \cdots, m. \tag{21}$$

If Y is finite dimensional, we get the following verification of the well-known fact that the Moore–Penrose inverse of any finite matrix exists and is unique.

COROLLARY 2. *If Y is finite dimensional, the BLSS of the problem $\mathscr{L}x = y$ exists and is unique for each $y \in Y$.*

Proof. For finite dimensional Y, $Y|^{Y^*} = Y^*$, or more crudely, $Y = Y^*$. Therefore, Corollary 5 in the Appendix shows that (21) is always satisfied.

Recall that functionals $f_1, \cdots, f_m$ in X^* were defined in the Introduction to be admissible constraints for L if and only if problem (1) (2) has a BLSS for every y in Y. Theorem 1 then states that $f_1, \cdots, f_m$ are admissible constraints for L iff they satisfy (20). A more explicit characterization is given in § 5.

We now turn to a characterization of the generalized inverse $\mathscr{L}^{\dagger}$ of $\mathscr{L}$, as defined in the Introduction.

COROLLARY 3. *The generalized inverse $\mathscr{L}^{\dagger}$ exists if and only if the constraints* (1) *satisfy condition* (20) *of Theorem* 1. *In this case, $\mathscr{L}^{\dagger}$ has the following properties.*

(22)
- (a) *$\mathscr{L}^{\dagger}$ is linear.*
- (b) $D(\mathscr{L}^{\dagger}) = Y$, $N(\mathscr{L}^{\dagger}) = Z$.
- (c) $R(\mathscr{L}^{\dagger}) = N(\mathscr{L})^{\perp} \cap F_0$.
- (d) $\mathscr{L}\mathscr{L}^{\dagger}\mathscr{L} = \mathscr{L}$, $\mathscr{L}^{\dagger}\mathscr{L}\mathscr{L}^{\dagger} = \mathscr{L}^{\dagger}$.
- (e) *$\mathscr{L}^{\dagger}\mathscr{L}$ is the orthogonal projector of F_0 onto $N(\mathscr{L})^{\perp}$.*
- (f) *$\mathscr{L}\mathscr{L}^{\dagger}$ is the orthogonal projector of Y onto $R(\mathscr{L})$.*

The proofs are all immediate.

4. Application. In this section, we apply Theorem 1 to the general two-point boundary value problem

$$Lx \equiv \frac{dx}{dt} + A(t)x = y, \tag{23}$$

$$Mx(a) + Nx(b) = 0, \tag{24}$$

where

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \qquad A = (a_{ij}),$$

the functions a_{ij} are assumed continuous on $[a, b]$, and M and N are $m \times n$ matrices with $0 \leqq m \leqq 2n$. We take

$$x \in X = C_n^1[a, b], \quad \text{and} \quad y \in Y = C_n[a, b]. \tag{25}$$

By well-known theorems, L defined by (23) maps *onto* Y. The problem (23) (24) does not have a solution in general; we ask whether it has a best least squares solution with respect to the usual L_2 norm

$$\|u\| = (u, u)^{1/2}, \quad \text{where} \quad (u, v) = \int_a^b v^*(t)u(t)\,dt. \tag{26}$$

Here $v^*(t)$ denotes the transpose of the vector $v(t)$. The boundary conditions (24) are to be thought of as linear functionals on X. Any such boundary conditions are just linear combinations of the $2n$ linear functionals defined by

$$\begin{aligned} f_i(x) &= x_i(a), \qquad i = 1, \cdots, n, \\ f_{n+i}(x) &= x_i(b), \qquad i = 1, \cdots, n. \end{aligned} \tag{27}$$

Note that the functionals (27) are unbounded in the topology induced by the norm (26). Clearly X and Y are not complete, and the mapping L in (23) is unbounded. These features are no obstacle to the application of Theorem 1. Previous studies [1], [8] and [11] of this problem have proceeded by completing Y to the Hilbert space $L_2[a, b]$ and extending X to the space of absolutely continuous functions with derivatives in $L_2[a, b]$. It is easy to show that the BLSS exists and is unique. However, in applications it is important to know the smoothness of a solution. An elaborate construction involving a generalized Green's matrix shows that if $y \in C_n[a, b]$ then $x^\dagger \in C_n^1[a, b]$. The present more general theory achieves this result directly.

According to Corollary 1, it is sufficient to show that $f_1, \cdots, f_{2n}$ satisfy (21). However it is clear that we could reorder the x_i's and interchange a and b, so that it suffices to consider f_1 of (27). In fact we establish in Lemma 3 a stronger result than (21). Note that in this example, X can be considered a subspace of Y, which as before is a subspace of Y^*.

LEMMA 3. *In example* (23)–(26) *the boundary functionals* (27) *satisfy*

$$f_i \in N(L)|^{X^*} \oplus L^*X, \qquad i = 1, \cdots, 2n. \tag{28}$$

Proof. By symmetry it suffices to consider f_1. Let $\Phi(t)$ denote the principal matrix solution defined by

$$\frac{d\Phi}{dt} + A(t)\Phi = 0, \qquad \Phi(a) = I. \tag{29}$$

Then the columns of Φ are a basis for $N(L)$; write them $\Phi^{(1)}, \cdots, \Phi^{(n)}$. The Gram matrix is

$$G=\int_a^b \Phi^*(t)\Phi(t)\,dt, \qquad \text{i.e.} \quad G_{ij}=(\Phi^{(i)}, \Phi^{(j)}).$$

Now for any $x \in X$,

$$f_1(x)=\langle\Phi, f_1\rangle G^{-1}(x, \Phi)+h_1(x), \tag{30}$$

where $h_1 \in N(L)^0=R(L^*)$, and

$$\begin{aligned}\langle\Phi, f_1\rangle &\equiv (\langle\Phi^{(1)}, f_1\rangle, \cdots, \langle\Phi^{(n)}, f_1\rangle)\\ &=(1, 0, \cdots, 0),\end{aligned}$$

$$(x, \Phi)\equiv \begin{bmatrix}(x, \Phi^{(1)})\\ \vdots \\ (x, \Phi^{(n)})\end{bmatrix}=\int_a^b \Phi^*(t)x(t)\,dt. \tag{31}$$

Hence,

$$\begin{aligned}h_1(x)&=x_1(a)-(1, 0, \cdots, 0)G^{-1}\int_a^b \Phi^*(t)x(t)\,dt\\ &=x_1(a)-\alpha_1,\end{aligned} \tag{32}$$

where α_1 is the first coefficient in the expansion

$$x=\sum_{i=1}^n \alpha_i\Phi^{(i)}+u, \qquad u \in N(L)^\perp.$$

Now we seek $z \in X$ such that

$$\begin{aligned}x_1(a)-\alpha_1=\langle x, h_1\rangle &= \langle x, L^*z\rangle\\ &=(Lx, z)\\ &=\int_a^b z^*(t)\left(\frac{dx}{dt}+Ax\right)dt\\ &=z^*x|_a^b+\int_a^b\left(-\frac{dz}{dt}+A^*z\right)^* x(t)\,dt,\end{aligned} \tag{33}$$

where we have integrated by parts, assuming $z \in X$. But (33) holds if z satisfies

$$z(b)=\begin{bmatrix}0\\0\\ \vdots\\0\end{bmatrix}, \qquad z(a)=-\begin{bmatrix}1\\0\\ \vdots\\0\end{bmatrix}, \quad \text{and} \tag{34}$$

$$-\frac{dz}{dt}+A^*z=-\Phi(t)G^{-1}\begin{bmatrix}1\\0\\ \vdots\\0\end{bmatrix}. \tag{35}$$

The general solution of (35) by the variation of parameters formula is, for arbitrary $\gamma \in \mathbf{R}^n$,

$$z(t)=\Phi^{*-1}(t)\left[\gamma+\int_a^t \Phi^*(s)\Phi(s)\,ds\,G^{-1}\begin{bmatrix}1\\0\\ \vdots\\0\end{bmatrix}\right].$$

Clearly the choice

$$\gamma=-\begin{bmatrix}1\\0\\ \vdots\\0\end{bmatrix}$$

satisfies both of the conditions in (34) simultaneously, hence $h_1=L^*z$ where

$$z(t)=-\Phi^{*-1}(t)\int_t^b \Phi^*(s)\Phi(s)\,ds\,G^{-1}\begin{bmatrix}1\\0\\ \vdots\\0\end{bmatrix} \in X.$$

From this lemma and Corollary 1 we immediately obtain

THEOREM 2. *The general two-point boundary value problem* (23)–(26) *has a unique BLSS* $x^\dagger \in C_n^1[a,b]$ *for any* $y \in C_n[a,b]$ *and any boundary conditions of the form* (24).

The applicability of this theory to more general boundary value problems is indicated, but will not be pursued here. Instances in which problems of the form (1) (2) arise are described in [6] and [12].

5. Explicit formulae. This section presents explicit formulae (in Theorems 3 and 4) which replace certain calculations involving $\mathscr{L}$ with calculations involving L and some matrix operations. The assumption is that it is easier to work with the unconstrained operator L than with the constrained operator $\mathscr{L}$. In the case of differential equations, this amounts to the observation that initial value problems are easier to solve than boundary value problems.

Let $\{\phi_1,\cdots,\phi_n\}$ be any basis for $N(L)$. (For ordinary differential equations, these are solutions of n initial value problems.) It is not assumed that these solutions have any special minimization or orthogonality properties. We use the same symbols ϕ_i to denote their images in $N(L)|^{X^*}$, and write

$$(x,\phi_i)=\langle x,\phi_i\rangle=\phi_i(x),\qquad x\in X,\quad i=1,\cdots,n.$$

Define the Gram matrix G as usual by

$$G_{ij}=(\phi_i,\phi_j),\qquad i,j=1,\cdots,n;$$

and use the given constraint functionals to define the $m\times n$ matrix $A=(a_{ij})$ by

$$a_{ij}=f_i(\phi_j)=\langle\phi_j,f_i\rangle,\qquad i=1,\cdots,m,\quad j=1,\cdots,n.$$

It follows from Corollary 5 in the Appendix that there exist unique constants $b_{ij}\in R$

and functionals $h_i \in R(L^*)$ such that

(36) $$f_i = \sum_{j=1}^{n} b_{ij}\phi_j + h_i, \qquad i = 1, \cdots, m.$$

Direct calculation shows that the matrices G, A and $B = (b_{ij})$ are related by

(37) $$B = AG^{-1}.$$

Introducing the notation

(38) $$\mathbf{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix}, \quad \boldsymbol{\phi} = \begin{bmatrix} \phi_1 \\ \vdots \\ \phi_n \end{bmatrix}, \quad \text{and} \quad \mathbf{h} = \begin{bmatrix} h_1 \\ \vdots \\ h_m \end{bmatrix},$$

we can write (36) as

(39) $$\mathbf{f} = AG^{-1}\boldsymbol{\phi} + \mathbf{h}.$$

Let A^* denote the transpose of the matrix A and $N(A)$, $R(A)$ denote the usual nullspace and range of A in $\mathbf{R}^n$ and $\mathbf{R}^m$ respectively. Suppose that the rank of A is r. Then we define an $n \times (n-r)$ matrix $C = (c_{ij})$ such that the columns of C are a basis for the nullspace of A. Similarly define the $m \times (m-r)$ matrix $E = (e_{ij})$, of which the columns are a basis for $N(A^*)$, and write

(40) $$e^{(i)} = \begin{bmatrix} e_{i1} \\ \vdots \\ e_{im} \end{bmatrix}, \qquad i = 1, \cdots, m-r.$$

It is convenient to extend the vector notation (38) further with the conventions

(41) $$\alpha^*\boldsymbol{\phi} = \sum_{i=1}^{n} \alpha_i\phi_i, \quad \text{where } \alpha \in \mathbf{R}^n,$$

(42) $$\mathbf{f}(x) = \langle x, \mathbf{f}\rangle = \begin{bmatrix} \langle x, f_1\rangle \\ \vdots \\ \langle x, f_m\rangle \end{bmatrix}, \quad \text{and} \quad E^*\mathbf{f} \equiv \begin{bmatrix} e^{(1)*}\mathbf{f} \\ \vdots \\ e^{(m-r)*}\mathbf{f} \end{bmatrix}.$$

LEMMA 4. *The space $F \cap N(L)^0 = L^*Z$ of Lemma 2 is spanned by the components of $E^*\mathbf{f}$. The space $Z = (LF_0)^0$ is spanned by the components of the solution $\boldsymbol{\psi} = (\psi_1, \cdots, \psi_{m-r})^*$ of*

(43) $$L^*\boldsymbol{\psi} = E^*\mathbf{f}.$$

If the given constraints $f_1, \cdots, f_m$ are linearly independent, then the components of $E^*\mathbf{f}$ and $\boldsymbol{\psi}$ are bases of their respective spans.

Proof. Clearly each component of $E^*\mathbf{f}$ is in F and annihilates every ϕ_j. Conversely, suppose $g \in F \cap N(L)^0$. Then

$$g = \sum_{i=1}^{m} \alpha_i f_i = \alpha^*\mathbf{f}, \quad \text{and for } j = 1, \cdots, n,$$

$$0 = g(\phi_j) = \alpha^* f(\phi_j) = \alpha^* A^{(j)},$$

which implies $\alpha \in N(A^*) = \text{span}(E)$, and so $g \in \text{span}(E^*\mathbf{f})$. The remaining assertions follow from Lemma 2 and the facts that L^* is one-to-one and E^* has full row rank.

In view of this lemma, the basic existence Theorem 1 can be rephrased as: The constraints (1) are admissible if and only if each component of $E^*\mathbf{f}$ lies in L^*Y. From this one easily obtains the following explicit characterization of a general set of admissible constraints.

THEOREM 3. *Given any positive integer m, any $m\times n$ real matrix A, any* $\mathbf{g}=(g_1,\cdots,g_n)^*$ *with* $g_i\in N(L)^0$, *and any* $\mathbf{x}=(x_1,\cdots,x_m)^*$ *with* $x_i\in X$, *define* $\mathbf{f}=(f_1,\cdots,f_m)^*$ *by*

$$\mathbf{f}=AG^{-1}\boldsymbol{\phi}+A\mathbf{g}+L^*L\mathbf{x}. \tag{44}$$

Then $f_1,\cdots,f_m$ *are admissible constraints for L. Furthermore, any set of admissible constraints for L can be written in this form* (44).

Henceforth we assume that the constraints (1) are admissible, and we consider the problem of actually calculating the BLSS of problem (1) (2). One approach is to use the "normal equation" which is well known in a Hilbert space context, see [2], [8], [9] and [10]. That is, $x\in F_0$ is a least squares solution (i.e. minimizes $\|x-y\|$) if and only if x satisfies

$$\mathscr{L}^*\mathscr{L}x=\mathscr{L}^*y. \tag{45}$$

If one is able to solve the normal equation (45) for x, it is then an easy matter to project x onto $N(\mathscr{L})^\perp$ and thus obtain $x^\dagger$. However, there are computational drawbacks associated with (45); the determination of $\mathscr{L}^*|_Y$ may be difficult and the product $\mathscr{L}^*\mathscr{L}$ is often ill-conditioned. Instead we proceed now to reduce the calculation of the BLSS to calculations involving finite matrices and the unconstrained operators L and L^*. Without any loss of generality we can assume that the constraints (1) are linearly independent.

THEOREM 4. *Suppose the constraints* (1) *are admissible and linearly independent. Then the BLSS* $x^\dagger$ *of problem* (1) (2) *for any* $y\in Y$ *is given by the following sequence of calculations*:

(46)
- (a) $\hat{y}=y-\boldsymbol{\psi}^*H^{-1}(y,\boldsymbol{\psi})$,
- (b) x_p *is any solution of* $Lx_p=\hat{y}$ (*not necessarily in* F_0),
- (c) $\alpha=-P_{T,S}A^\dagger\langle x_p,\mathbf{f}\rangle-P_{S,T}G^{-1}(x_p,\boldsymbol{\phi})$,
- (d) $x^\dagger=\alpha^*\boldsymbol{\phi}+x_p$.

The new symbols introduced in these formulae are defined as follows:

$H=(\psi_i,\psi_j)$,
$G=(\phi_i,\phi_j)$ (*G and H are Gram matrices*),
$A^\dagger=$ *Moore–Penrose inverse of matrix A*,
$P_{T,S}=n\times n$ *idempotent matrix* (*projector*) *with range T and nullspace S*,
$P_{S,T}=I_n-P_{T,S}$ ($I_n=n\times n$ *identity matrix*),
$T=$ *range of matrix* $G^{-1}A^*$,
$S=$ *nullspace of A*.

Remarks. (i) The ψ_i's are defined by

$$L^*\psi_i=e^{(i)*}\mathbf{f},\qquad i=1,\cdots,m-r; \tag{47}$$

see (40) through (43). Equations (47) are not as formidable as they might at first appear; L^* is often very similar in form to L, and Theorem 1 guarantees that the ψ_i's are not general functionals in Y^* but actually lie in the more familiar space Y. Thus (47) may be no harder to solve than (2). (See § 4.)

(ii) The invertibility of H follows from the assumed linear independence of the constraints, and Lemma 4.

(iii) It is not necessary to use the Moore–Penrose inverse $A^\dagger$ in (c), in fact, in the notation of [2], any {1}-inverse will do, that is any matrix $A^{(1)}$ satisfying the first Penrose condition

$$AA^{(1)}A = A. \tag{48}$$

Recent advances in numerical linear algebra (e.g. [4]) make the computation of $A^\dagger$ or $A^{(1)}$ a routine matter.

Proof of Theorem 4. The preceding theorems guarantee that any $y \in Y$ has a unique projection on $R(\mathscr{L}) = LF_0$, and it is given by (46a). Now the BLSS of problem (1) (2) is the same as the minimum norm solution of problem (1) (46b). This must certainly be of the form (46d) for some α, so it remains only to derive (46c).

Apply the constraints (1) to (d) and find that α satisfies

$$A\alpha = -\langle x_p, \mathbf{f}\rangle, \tag{49}$$

where the right hand side is automatically in the range of A. Define

$$\xi = -A^\dagger\langle x_p, \mathbf{f}\rangle; \tag{50}$$

then the general solution of (49) is

$$\alpha = \xi + \eta, \qquad \eta \in N(A). \tag{51}$$

Let C be an $n \times (n-r)$ matrix whose columns form a basis of $N(A)$. Then $\eta = C\sigma$ for some $\sigma \in \mathbf{R}^{n-r}$, and a basis for $N(\mathscr{L})$ is given by the components of $C^*\boldsymbol{\phi}$. Now η is determined by the condition that $x^\dagger \in N(\mathscr{L})^\perp$, i.e.

$$\begin{aligned} 0 &= (\boldsymbol{\phi}^*\alpha + x_p, C^*\boldsymbol{\phi}) \\ &= C^*G\xi + C^*GC\sigma + C^*(x_p, \boldsymbol{\phi}). \end{aligned}$$

Hence, since C^*GC is nonsingular

$$\begin{aligned} \eta &= C\sigma \\ &= -C(C^*GC)^{-1}C^*[G\xi + (x_p, \boldsymbol{\phi})]. \end{aligned} \tag{52}$$

Since G is positive definite and Hermitian, it has a unique positive definite Hermitian square root K. Employing the reverse order property established by Greville [5], we can write

$$(C^*GC)^{-1} = (C^*KKC)^{-1} = (KC)^\dagger(C^*K)^\dagger.$$

Therefore

$$\begin{aligned} C(C^*GC)^{-1}C^*G &= C(KC)^\dagger(C^*K)^\dagger C^*KK \\ &= K^{-1}KC(KC)^\dagger P_{R(KC)}K \\ &= K^{-1}P_{R(KC)}K \\ &= P_{S,T}, \end{aligned} \tag{53}$$

where $P_{R(KC)}$ is the orthogonal projector on the range of KC. Substituting (50), (52) and (53) into (51) gives (46c).

Theorem 4 has been applied in [7] to construct a shooting algorithm for the BLSS of the general linear two-point boundary value problem in § 4.

Appendix. A basic tool in this paper is the projection theorem for a pre-Hilbert space, which we use in the following form, only slightly modified from that in [9] or [13].

THEOREM 5 (Projection Theorem). *Let V be an inner product space and let M be a subspace of V. Given any $v \in V$, a necessary and sufficient condition that $m_0 \in M$ be a unique minimizing vector satisfying*

$$\|v - m_0\| \leqq \|v - m\| \quad \text{for all } m \in M,$$

is that $v - m_0$ be orthogonal to M, that is

$$(v - m_0, m) = 0 \quad \text{for all } m \in M.$$

If m_0 exists, it is unique. If M is complete, m_0 exists.

An important consequence of the projection theorem is:

COROLLARY 4. *If M is a complete subspace of an inner product space V, then*

$$V = M \oplus M^{\perp}.$$

Proof. See [9].

LEMMA 5. *If S is a subset of a linear space X, then*

$$[S] = (S^0)_0. \tag{54}$$

For a proof based on Zorn's lemma, see Theorem 1.9A in [13]. Note that if consideration is restricted to bounded linear functionals and we denote the *normed annihilator* of S by

$$S^{0\prime} = S^0 \cap X',$$

then (54) becomes

$$[S] \subset (S^{0\prime})_0 = \overline{[S]},$$

where the bar indicates closure [13].

LEMMA 6. *If F is a finite dimensional subspace of a dual space X^*, then*

$$F = (F_0)^0. \tag{55}$$

Proof. Clearly $F \subset (F_0)^0$. For the reverse inclusion, construct a basis $\{f_1, \cdots, f_m\}$ of F and a set $\{x_1, \cdots, x_m\}$ in X satisfying

$$\langle x_i, f_j \rangle = \delta_{ij}, \qquad i, j = 1, \cdots, m.$$

This can be done using a double Gram–Schmidt procedure, by Theorem 1.71B of [13]. Any $x \in X$ can be written

$$x = \left(\sum_{j=1}^{m} x_j \langle x, f_j \rangle \right) + \left(x - \sum_{j=1}^{m} x_j \langle x, f_j \rangle \right),$$

from which it easily follows that

$$X = [x_1, \cdots, x_m] \oplus F_0.$$

Now suppose $f^* \in (F_0)^0$. Then for any $x \in X$,

$$f^*\left(x - \sum_{j=1}^{m} x_j \langle x, f_j \rangle \right) = 0,$$

so

$$f^*(x) = \sum_{j=1}^{m} f^*(x_j) f_j(x),$$

which shows $f^* \in F$.

For a general subspace F of X^*, equation (55) must be weakened to

$$F \subseteq (F_0)^0.$$

This is true even if the functionals are restricted to X'.

LEMMA 7.

$$\text{(56)} \qquad \begin{aligned} &\text{(a)} \quad R(\mathcal{L})^0 = N(\mathcal{L}^*), \\ &\text{(b)} \quad R(\mathcal{L}) = N(\mathcal{L}^*)_0. \end{aligned}$$

Proof. Part (a) follows directly from (8), the definition of $\mathcal{L}^*$. Part (b) is a consequence of Lemma 5.

Note that $N(\mathcal{L}^*)$ is precisely the subspace Z of §§ 3 and 5. For the normed dual, (b) is no longer valid.

LEMMA 8.

$$\text{(57)} \qquad \begin{aligned} &\text{(a)} \quad N(L) = R(L^*)_0, \\ &\text{(b)} \quad N(L)^0 = R(L^*). \end{aligned}$$

Proof. Part (a) follows from (7), the definition of L^*, (see Theorem 1.91D of [13]). In part (b), $R(L^*) \subseteq N(L)^0$ follows from (7), but the reverse inclusion is false for general L. However, in our case, $N(L)$ is finite dimensional, so by the corollary to the projection theorem,

$$\text{(58)} \qquad X = N(L) \oplus N(L)^\perp,$$

i.e. for any $x \in X$ there exist unique $x_0 \in N(L)$ and $x_1 \in N(L)^\perp$ such that

$$\text{(59)} \qquad x = x_0 + x_1.$$

Now the restriction of mapping L to $N(L)^\perp$ is one-to-one and onto; for any $y \in Y$ there exists a unique $x_1 \in N(L)^\perp$ such that $Lx_1 = y$. To show $N(L)^0 \subseteq R(L^*)$, take any $x^* \in N(L)^0$. Define a corresponding $y^* \in Y^*$ by

$$\langle y, y^* \rangle = \langle x_1, x^* \rangle$$

where x_1 is the unique solution in $N(L)^\perp$ of $Lx_1 - y$. Now for any $x \in X$,

$$\langle x, L^* y^* \rangle = \langle Lx, y^* \rangle = \langle Lx_1, y^* \rangle = \langle x_1, x^* \rangle$$

where x_1 is as defined in (59). But $x^* \in N(L)^0$ implies $\langle x, x^* \rangle = \langle x_1, x^* \rangle$ for all $x \in X$, so we conclude $x^* = L^* y^*$.

COROLLARY 5.

$$X^* = N(L)|^{X^*} \oplus R(L^*).$$

This follows from Lemma 8 by an argument similar to that in the proof of Lemma 6.

REFERENCES

[1] J. S. BRADLEY, *Generalized Green's matrices for compatible differential systems*, Michigan Math. J., 13 (1966), pp. 97–108.

[2] A. BEN-ISRAEL AND T. N. E. GREVILLE, *Generalized Inverses: Theory and Applications*, Wiley-Interscience, New York, 1974.

[3] S. GOLDBERG, *Unbounded Linear Operators*, McGraw-Hill, New York, 1966.

[4] G. H. GOLUB AND C. REINSCH, *Singular value decomposition and least squares solutions*, Numer. Math., 14 (1970), pp. 403–420.

[5] T. N. E. GREVILLE, *Note on the generalized inverse of a matrix product*, SIAM Rev., 8 (1966), pp. 518–521.

[6] H. B. KELLER AND W. F. LANGFORD, *Iterations, perturbations and multiplicities for nonlinear bifurcation problems*, Arch. Rational Mech. Anal., 48 (1972), pp. 83–108.

[7] W. F. LANGFORD, *A shooting algorithm for the best least squares solution of two-point boundary value problems*, SIAM J. Numer. Anal., 14 (1977), pp. 527–542.

[8] W. S. LOUD, *Some examples of generalized Green's functions and generalized Green's matrices*, SIAM Rev., 12 (1970), pp. 194–210.

[9] D. G. LUENBERGER, *Optimization by Vector Space Methods*, John Wiley, New York, 1969.

[10] M. Z. NASHED, *Generalized inverses, normal solvability and iteration for singular operator equations*, Nonlinear Functional Analysis and Applications, L. B. Rall ed., Academic Press, New York, 1971.

[11] W. T. REID, *Generalized Green's matrices for compatible systems of differential equations*, Amer. J. Math., 53 (1931), pp. 443–459.

[12] I. STAKGOLD, *Branching of solutions of nonlinear equations*, SIAM Rev., 13 (1971), pp. 289–332.

[13] A. E. TAYLOR, *Introduction to Functional Analysis*, John Wiley, New York, 1958.

[14] S. ZLOBEC, *On computing the best least squares solutions in Hilbert space*, Rend. Circ. Mat. Palmero, 25 (1976), pp. 1–15.

SIAM. J. MATH. ANAL.
Vol. 9, No. 6, December 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0009 $01.00/0

ANALYTIC CONTINUATION VIA HADAMARD'S PRODUCT*

BARBARA FROMM CHAMBERS†

Abstract. This paper presents an operational procedure derived from Hadamard's convolution product which is used to construct continuations of analytic functions in the form of integral functional representations. These representations are more useful in the study of analytic properties than the underlying Taylor's series, and the method extends the previously well-established continuation results of Borel and Mittag–Leffler.

1. Introduction. The purpose of this research note is to develop a generalized procedure for constructing continuations of analytic functions which may exist beyond any of the Taylor series' circles of convergence. The method is based on Hadamard's product and includes, as special cases, the established results of Borel [3, p. 122] and Mittag–Leffler [9, pp. 431–438].

To begin, consider a function, $f(z)$, analytic in a region $\mathbf{R}$. This function, as an abstract equivalence class of mappings, is characterized by its representations, i.e., mathematical expressions which coincide with $f(z)$ in their individual regions of analyticity. Typical representations are tables, graphs, figures, Taylor series, Laurent series, Weierstrass factorizations, Mittag–Leffler expansions, Laplace transforms, differential equation, integral formulae, etc. The properties of any such function can then be investigated by means of these representations, and these properties can then be extended to as large a subset of $\mathbf{R}$ as permitted by analytic continuation.

Of course, $f(z)$ can be represented by a Taylor series in the neighborhood of every point in $\mathbf{R}$, and $f(z)$ can be completely characterized by the collection of Taylor series whose circles of convergence cover $\mathbf{R}$. However, this is usually a large and unwieldy collection from which it is difficult to derive functional properties. The following problem is posed in this paper: Given one of the Taylor series in this collection, which may be taken to be about the origin, construct a single representation of $f(z)$ which contains the given series' circle of convergence and also points in other circles of convergence of Taylor series which may continue the chosen one.

The procedure for solving this problem involves expanding radially through arcs of the circle of convergence which do not contain barrier points. Thus the continuation will be confined to the principal star, i.e. the largest open subset, $\mathbf{P}$, of $\mathbf{R}$ such that if a point $p \in \mathbf{P}$ then the radius segment $\mathbf{op} \subset \mathbf{P}$. ($\mathbf{P}$ does not necessarily coincide with $\mathbf{R}$.) Any such continuation would necessarily exist in a subset of $\mathbf{P}$, and is called a radial continuation. Finally, the techniques developed in this paper have also been used by the author [4, pp. 77–80], to study the genus, order and singularities of products and continuations, and [5] to generate a transform calculus which is used to study stress system applications.

2. The continuation. First, consider Hadamard's product.

DEFINITION. Let

$$f(z)=\sum a_n z^n, \qquad |z|<R_1,$$

$$g(z)=\sum b_n z^n, \qquad |z|<R_2;$$

then the series $h(z)=\sum a_n b_n z^n$ is called the Hadamard product and the function h is denoted $f*g$. Hadamard's theorem [4, p. 75] states that this series converges for

* Received by the editors November 18, 1975, and in final revised form April 11, 1977.
† Department of Mathematics, George Mason University, Fairfax, Virginia 22030.

$|z|<R$, where R is at least equal to the product R_1R_2. Also, for $|z|<R_1R_2$, $|\zeta|<R_1|z/\zeta|<R_2$ it is well known that

$$h(z)=\frac{1}{2\pi i}\int_{\mathscr{C}} f(\zeta)g(z/\zeta)\,d\zeta/\zeta,$$

where $\mathscr{C}$ is a simple contour encircling the origin. This form displays the convolution quality of the product, and indeed provides a possible continuation of $h(z)$ by deforming $\mathscr{C}$. Further details about this particular result are presented in [4].

The next step in the continuation of the given function $f(z)$ is the selection of an integral function, $g(z)$, to be used as a test weighting function to represent $f(z)$ over the entire complex plane. It is based on the following lemma concerning the Hadamard product of an analytic function and an integral function.

LEMMA. *Let* $f(z)=\sum a_n z^n$, $|z|<R_1$, *and* $g(z)$ *be integral, i.e.,* $g(z)=\sum b_n z^n$, *for all* z; *then* $h(z)=f*g(z)=\sum a_n b_n z^n$ *is an integral function.*

Proof. By Hadamard's theorem, we may assume that $\sum a_n b_n z^n$ converges for $|z|<R$, where $R>0$. Hence $1/R\geqq 0$. But it is known that

$$\begin{aligned}1/R&=\limsup_{n\to\infty}|a_n b_n|^{1/n}\\&\leqq \limsup_{n\to\infty}|a_n|^{1/n}\limsup_{n\to\infty}|b_n|^{1/n}\\&\leqq (1/R_1)\limsup_{n\to\infty}|b_n|^{1/n}.\end{aligned}$$

But since $g(z)$ is an integral function, $\limsup_{n\to\infty}|b_n|^{1/n}=0$. Therefore, $0\leqq 1/R\leqq 0$, and $\sum a_n b_n z^n$ converges for all z.

Of course, $h(z)$ is not an analytic continuation of $f(z)$, but rather its weighted average over the integral function $g(z)$ which spreads the analytic existence of $f(z)$ within $\mathbf{R}$ over the entire complex plane. The final step is to condense $h(z)$ (viewed as a distorted $f(z)$), back into a smaller domain in such a way so as to recover $f(z)$, in a different form, which would represent a continuation of $f(z)$ past any arc in $|z|=R_1$ through which a continuation existed.

Proceeding formally, denote a test integral function by $\Psi(z)=\sum \varphi_n z^n$, where none of the φ_n are zero. Then form the product integral function, $f*\Psi(z)$. We now propose to find a linear functional, $\mathscr{L}$, operating on the class of complex valued functions and a parametric function, $\rho(z)$, such that an analytic continuation, $f_c(z)$ of $f(z)$ can be expressed as follows in a continuation region R_c containing $\{z||z|<R_1\}$:

$$f_c(z)=\mathscr{L}_\zeta[f*\Psi(z\rho(\zeta))],\qquad z\in R_c.$$

Here z is held constant, and $\mathscr{L}$ operates through the argument, ζ, of the parametric function ρ. Since $\mathscr{L}$ is linear, and $f*\Psi$ is integral in its argument, we have for our continuation, $f_c(z)$:

$$\begin{aligned}f_c(z)&=\mathscr{L}_\zeta[f*\Psi(z\rho(\zeta))]\\&=\mathscr{L}_\zeta[\textstyle\sum a_n\varphi_n z^n\rho^n(\zeta)]\\&=\textstyle\sum a_n\varphi_n\mathscr{L}_\zeta[\rho^n(\zeta)]z^n,\end{aligned}$$

which must coincide with $f(z)=\sum a_n z^n$, within the region $|z|<R_1$.

If such an $\mathscr{L}$ and ρ exist, then necessarily they satisfy the following conditions:

$$\mathscr{L}_\zeta[\rho^n(\zeta)]=1/\varphi_n,\qquad n=0,1,2,\cdots.$$

(Note: the requirement that the test function, $\Psi(z)$, have every term in its Maclaurin's series is necessary here. This is a reasonable requirement, since otherwise some information about $f(z)$ within $|z|<R_1$, i.e. some of its coefficients, would be lost in the associated product integral function, $f*\Psi(z)$, and on condensing back through $\mathscr{L}$ and ρ we would not in general be able to obtain a continuation of $f(z)$.) We now apply this general formulation to some specific continuations.

3. Illustrative methods.

3.1. The Borel method. Choose the test function, $\Psi(z)=\exp(z)=\sum\{1/\Gamma(1+n)\}z^n$. Clearly, $1/\Gamma(1+n)\neq 0$ for $n=0,1,2,\cdots$, and $\Psi(z)$ is an acceptable test function. The basic conditions given above become:

$$\mathscr{L}_\zeta[\rho^n(\zeta)]=\Gamma(1+n)=\int_0^\infty e^{-\zeta}\zeta^n\,d\zeta,\qquad n=0,1,2,\cdots.$$

By choosing $\rho(\zeta)=\zeta$ and $\mathscr{L}[F]=\int_0^\infty e^{-\zeta}F(\zeta)\,d\zeta$, we then have

$$f_c(z)=\mathscr{L}_\zeta[f*\Psi(z\rho(\zeta))\,d\zeta]=\int_0^\infty e^{-\zeta}f*\exp(z\zeta)\,d\zeta.$$

By expressing $f*\exp(z)$ in the contour integral form given in the introduction, Borel [3, p. 122] has shown that $f_c(z)$ is an analytic continuation of $f(z)$ into a polygonal region R_c containing $\{z\,|\,|z|<R_1\}$.

This well-known polygon is tangent to the original circle of convergence of $f(z)$ at the barrier point(s) on the circle. Depending on the number and location of the barrier points of $f(z)$, R_c may be a semi-infinite region, but if a principal star continuation region exists, then R_c will not extend into the star beyond the polygonal edges. Mittag–Leffler presented a refinement of Borel's approach which extended the area of continuation R_c further into the principal star. We will now consider how this refinement is obtained from our method.

3.2. The Mittag–Leffler method. This method is based on the test function $\Psi(z)=E_\alpha(z)=\sum\{1/\Gamma(1+\alpha n)\}z^n$ which is an entire function for all $\alpha>0$. Since $E_1(z)=e^z$, this method should include the Borel method as a special case. However, note also that as $\alpha\to 0$, $E_\alpha(z)\to 1/(1-z)$. This is not an integral function, but does have only one isolated singularity at $z=1$. Functionally, it is also the Cauchy integrand for $z_0=1$, and, as we shall see later, this is very helpful in examining continuations into the entire principal star.

In this case, the basic conditions become:

$$\mathscr{L}_\zeta[\rho^n(\zeta)]=\Gamma(1+\alpha n)=\int_0^\infty e^{-\zeta}\zeta^{\alpha n}\,d\zeta,\qquad n=0,1,2,\cdots.$$

By choosing here $\rho(\zeta)=\zeta^\alpha$ and $\mathscr{L}$ as in the Borel method we then have

$$f_c(z)=\int_0^\infty e^{-\zeta}f*E_\alpha(z\zeta^\alpha)\,d\zeta.$$

Mittag-Leffler [9, p. 434] has shown that $f_c(z)$ is an analytic continuation of $f(z)$ into a region R_c which we will now examine.

Without loss of generality, we may consider one of the isolated singularities of $f(z)$ to be located at $z=1$. The principal star associated with this singularity will be a region in the right half plane cut from $+1$ to $+\infty$. Now, the boundary of the continuation region will then depend on the value of the parameter α. For $\alpha = 1$, it is easy to see from the integral for $f_c(z)$ that R_c becomes the Borel polygon. At $z = 1$ the boundary is a vertical line with that portion of R_c satisfying $\mathrm{Re}\,(z) < 1$. For $\alpha > 1$, this line folds towards the origin following a parabola-like shape, and in general R_c is a curved, convex region contained in the Borel polygon. However, for $\alpha < 1$, the line folds outward around the positive real axis, with the region of convergence being to the left. Again it is a conic-like shape, and in fact for $\alpha = 1/2$ it becomes exactly one branch of the hyperbola, $x^2 - y^2 = 1$. These variations are shown in Fig. 1.

For other barrier points, similar outward folding curves form the boundary and have as their axes the radial lines through the barrier points. In general, Hille [7, p. 69] has shown that the cusp-shaped region formed by the intersection of these conic-like edges is the region of convergence of the Mittag–Leffler method, and extends further into the principal star than the Borel polygon. Finally, as $\alpha \to 0$, the arms of the cusps fold out along the radial lines, and for α chosen small enough, any point in the interior of the principal star can be made to lie within the Mittag–Leffler region R_c. We will now consider other new methods and examine their convergence within the principal star.

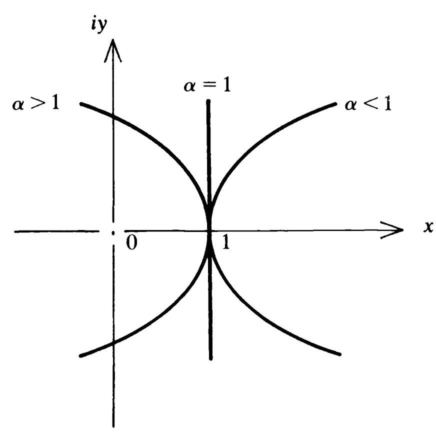

FIG. 1. *Mittag–Leffer continuation*

3.3. Other methods. The next method is based on the use of LeRoy's sum as the test function, i.e.,

$$\Psi(z) = R_\alpha(z) = \sum \{\Gamma(1+\alpha n)/\Gamma(1+n)\} z^n.$$

For $0 \leqq \alpha < 1$, this is an integral function, since if R is its radius of convergence, we have that

$$1/R = \limsup_{n\to\infty} \{\Gamma(1+\alpha n)/\Gamma(1+n)\} = \limsup_{n\to\infty} K n^{\alpha-1} = 0.$$

The basic conditions become here:

$$\mathscr{L}[\rho^n] = \Gamma(1+n)/\Gamma(1+\alpha n), \qquad n = 0, 1, 2, \cdots.$$

From Magnus and Oberhettinger [8, p. 9], we can write,

$$1/B(x, y) = \frac{(x+y-1)}{2\pi i} \int_{1/2-i\infty}^{1/2+i\infty} \sigma^{-x}(1-\sigma)^{-y}\, d\sigma$$

where $\mathrm{Re}(x+y)>1$, and we have chosen $a=b=1/2$ and $\sigma=1/2+it$ in the given formula. Using the fact that

$$B(x, y)=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)},$$

and letting $x=1+\alpha n$, $y=(1-\alpha)n$, $x+y=1+n$, we have

$$\frac{\Gamma(1+n)}{\Gamma(1+\alpha n)}=\frac{\Gamma((1-\alpha)n)}{B(1+\alpha n,(1-\alpha)n)}=\frac{n\Gamma((1-\alpha)n)}{2\pi i}\int_{1/2-i\infty}^{1/2+i\infty}\sigma^{-1-\alpha n}(1-\sigma)^{-(1-\alpha)n}\,d\sigma.$$

But since $z\Gamma(z)=\Gamma(z+1)$, we have

$$n\Gamma((1-\alpha)n)=\frac{\Gamma(1+(1-\alpha)n)}{(1-\alpha)}=\frac{1}{1-\alpha}\int_0^\infty e^{-\tau}\tau^{(1-\alpha)n}\,d\tau.$$

Combining, we finally obtain for the LeRoy coefficients,

$$\frac{\Gamma(1+n)}{\Gamma(1+\alpha n)}=\frac{1}{2\pi i(1-\alpha)}\int_0^\infty d\tau\int_{1/2-i\infty}^{1/2+i\infty}e^{-\tau}\left[\frac{\tau^{(1-\alpha)}}{\sigma^\alpha(1-\sigma)^{(1-\alpha)}}\right]^n\frac{d\sigma}{\sigma}.$$

By choosing $\rho(\zeta)=\zeta$ and

$$\mathscr{L}[F]=\frac{1}{2\pi i(1-\alpha)}\int_0^\infty d\tau\int_{1/2-1\infty}^{1/2+i\infty}e^{-\tau}F\left[\frac{z\tau^{(1-\alpha)}}{\sigma^\alpha(1-\sigma)^{(1-\alpha)}}\right]\frac{d\sigma}{\sigma},$$

we have for the continuation

$$f_c(z)=\frac{1}{2\pi i(1-\alpha)}\int_0^\infty d\tau\int_{1/2-i\infty}^{1/2+i\infty}e^{-\tau}f*R_\alpha\left[\frac{z\tau^{(1-\alpha)}}{\sigma^\alpha(1-\sigma)^{(1-\alpha)}}\right]\frac{d\sigma}{\sigma}.$$

To demonstrate that this actually is an analytic continuation, we observe that for finite τ, $\sigma, f*R_\alpha$ has a uniformly convergent series expansion in z, clearly given by

$$f*R_\alpha(z)=\sum\frac{\Gamma(1+\alpha n)}{\Gamma(1+n)}a_nz^n.$$

By choosing $0\leqq\tau\leqq\tau_0$, $-\omega_0\leqq I_m(\sigma)\leqq\omega_0$, we can interchange summation and integration over this rectangle. Setting $|z|<R_1$, to obtain uniform convergence in τ_0 and ω_0 and then letting these bounds become infinite, we finally obtain $f_c(z)=f(z)$ for $|z|<R_1$. Of course, this demonstrates only that theoretically $f_c(z)$ is an analytic continuation. If $f(z)$ can be continued beyond its circle of convergence, this result provides no further information. We would now like to examine the behavior of $f_c(z)$ in the principal star.

First note that for $\alpha=0$, $R_0(z)=e^z$, which generated the Borel method. Thus we should obtain continuations at least into the Borel polygon outside of the circle of convergence. Next, because of the functional linearity of these methods and the synthesizing structure of the Cauchy integral formula, Hille [7, p. 69] has shown that in general it is sufficient to examine the extent of continuation obtained by using the geometric series

$$f(z)=1/(1-z)=\sum z^n, \qquad |z|<1.$$

Clearly the only barrier point is $z=1$, and $a_n=1$, $n=0,1,2\cdots$. The principal star for $f(z)$ is the complex plane cut from $+1$ to $+\infty$. Because $[1/(1-z)]*R_\alpha(z)=R_\alpha(z)$, we have

$$f_c(z)=\frac{1}{2\pi i(1-\alpha)}\int_0^\infty d\tau\int_{1/2-i\infty}^{1/2+i\infty} e^{-\tau}R_\alpha\left[\frac{z\tau^{(1-\alpha)}}{\sigma^\alpha(1-\sigma)^{(1-\alpha)}}\right]\frac{d\sigma}{\sigma}.$$

To determine the region of convergence of this linear transform, we consider first the asymptotic behavior of $R_\alpha(z)$. Barnes [1, p. 283] has shown that

$$R_\alpha(z)\sim\left(\frac{2\pi}{1-\alpha}\right)^{1/2}(\alpha z)^{1/(2(1-\alpha))}\exp\left[(1-\alpha)\alpha^{\alpha/(1-\alpha)}z^{1/(1-\alpha)}\right],$$

$$|\arg(z)|<(1-\alpha)\frac{\pi}{2},$$

and elsewhere tends to zero. For large values of τ, the integrand in $f_c(z)$ thus behaves like

$$\left[\frac{\alpha^{1/(1-\alpha)}z^{1/(1-\alpha)}\tau}{2\pi(1-\alpha)(1-\sigma)\sigma^{(2-\alpha)/(1-\alpha)}}\right]^{1/2}\exp\left[-\tau\left(1-\frac{(1-\alpha)\alpha^{\alpha/(1-\alpha)}z^{1/(1-\alpha)}}{\sigma^{\alpha/(1-\alpha)}(1-\sigma)}\right)\right].$$

Because both σ and $1-\sigma$ are bounded away from the origin on their path of integration, the transform will converge in the region in which z is restricted by

$$\operatorname{Re}\left[1-\frac{(1-\alpha)\alpha^{\alpha/(1-\alpha)}z^{1/(1-\alpha)}}{\sigma^{\alpha/(1-\alpha)}(1-\sigma)}\right]>0,$$

where $1/2-i\infty<\sigma<1/2+i\infty$.

Introducing the parameter ω by $\sigma=1/2+i\omega/2, -\infty<\omega<\infty$, letting $z=re^{i\theta}$, $|\theta|<(1-\alpha)\pi/2$, taking real parts and rearranging, we have

$$r\cos^{(1-\alpha)}\left[\frac{1}{1-\alpha}(\theta-(2\alpha-1)\phi)\right]\cos\phi<\frac{1}{2\alpha^\alpha(1-\alpha)^{(1-\alpha)}}$$

where $\phi=\tan^{-1}(\omega)$.

For the argument θ in its permissible range, there will be a value $\phi_0(\theta,\alpha)$ for which the left hand side of this relation is a maximum, which is given by

$$(2\alpha-1)\tan\left[\frac{1}{1-\alpha}(\theta-(2\alpha-1)\phi_0)\right]=\tan(\phi_0).$$

This then fixes a range for r within the region of convergence. Because of the dominating effect of the ϕ-modulus factor over the ϕ-phase factor, ϕ_0 is bounded and in general $|\phi_0|<\pi/4$.

Let us determine this region explicitly for $\alpha=1/2$. From the above we have $\phi_0=0$, $|\theta|<\pi/4$, and the boundary becomes

$$r\cos^{1/2}(2\theta)<1.$$

The boundary of this region is the right branch of a hyperbola about the x-axis, asymptotic to $\theta=\pm\pi/4$, and passing through $z=1$. This region is shown in Fig. 2 and extends into the principal star beyond the Borel polygon.

Finally, the asymptotic expansion chosen above is not particularly convenient for studying the behavior of the continuation as $\alpha\to0$ or $\alpha\to1$. We know that as $\alpha\to0$, $R_\alpha(z)\to e^z$, uniformly in z and as $\alpha\to1$, $R_\alpha(z)\to1/(1-z)$. By choosing asymptotic expansions for $R_\alpha(z)$ compatible with these elementary functional limits it is possible to show first that as $\alpha\to0$, the continuation region becomes the Borel polygon, and

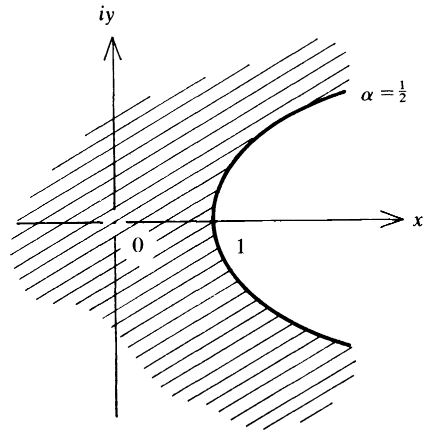

FIG. 2 *LeRoy continuation.*

second that as $\alpha \to 1$ the region extends as far into the principal star as desired. This latter property will be presented in more detail in the last example below.

For a general function, similar regions exist for each barrier point and, as noted by Hille [7, p. 69], the total continuation region is the intersection of all of them. If there are a finite number of barrier points, the continuation region R_c will be a cusp-shaped figure extending from the edges of the Borel polygon and filling out into the principal star as $\alpha \to 1$.

For the next example, choose from Bateman [2, p. 211]. the integral function

$$G_\beta(z;\lambda)=\sum \frac{1}{(n+\lambda)^\beta \Gamma(n+1)} z^n, \qquad |\arg (n+\theta)^\beta| < \pi,$$

β a nonnegative integer, and λ not zero or a negative integer. By using the relation $z\Gamma(z)=\Gamma(z+1)$ recursively, we can write

$$(n+\lambda)^\beta \Gamma(n+1)=\int_0^\infty e^{-\zeta}\zeta^n P_\beta(\zeta,\lambda)\, d\zeta$$

where $P_\beta(\zeta,\lambda)$ is a polynomial in ζ of degree β. If we denote $P_\beta(\zeta,\lambda)=\sum^\beta a_k(\zeta,\lambda)\zeta^k$, then the coefficients a_k are given by

$$a_0(0,\lambda)=1,$$

$$a_0(1,\lambda)=(\lambda-1), \quad a_1(1,\lambda)=1,$$

$$a_0(2,\lambda)=(\lambda-1)^2, \quad a_1(2,\lambda)=(2\lambda-3), \quad a_2(2,\lambda)=1,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$a_0(p,\lambda)=(\lambda-1)^p, \quad a_r(p,\lambda)=(\lambda-r-1)a_r(p-1,\lambda)+a_{r-1}(p-1,\lambda),$$

$$r=1,2,\cdots,p-1, \qquad a_p(p,\lambda)=1.$$

From the obvious choice of ρ and L we obtain the continuation as

$$f_c(z)=\int_0^\infty e^{-\zeta} P_\beta(\zeta,\lambda) f*G_\beta(z\zeta;\lambda)\, d\zeta.$$

To examine the region of convergence, set $f(z)=1/(1-z)$. Then Barnes [1, p. 265] has shown that for Re $(z)>0$ and $|z|$ large

$$G_\beta(z;\lambda)\sim e^z/z^\beta$$

and elsewhere tends to zero. Thus the integrand of $f_c(z)$ behaves like

$$\exp\left[-\zeta(1-z)\right]\frac{P_\beta(\zeta,\lambda)}{z^\beta\zeta^\beta}$$

and for $\mathrm{Re}(1-z)>0$, the integral will converge and the continuation will exist. This region is the half plane left of the vertical line passing through $z=1$, and is just the Borel polygon for $1/(1-z)$. Similar convergence would exist at other singularities of a general function $f(z)$, and this formula would yield continuations into the Borel polygon only.

For the final example, again choose from Bateman [2, p. 211] the integral function,

$$E_\alpha(z;\lambda,\beta)=\sum\frac{1}{(n+\lambda)^\beta\Gamma(1+\alpha n)}z^n,\qquad 0<\alpha<2,$$

β a nonnegative integer, and λ not zero or a negative integer.

By slight rearrangement of the relationships in the last example, we may write

$$(n+\lambda)^\beta\Gamma(1+\alpha n)=\frac{1}{\alpha^\beta}\int_0^\infty e^{-\zeta}\zeta^{\alpha n}P_\beta(\zeta,\alpha\lambda)\,d\zeta$$

where P is the polynomial of degree β given earlier. The method produces the following continuation formula,

$$f_c(z)=\frac{1}{\alpha^\beta}\int_0^\infty e^{-\zeta}\zeta^{\alpha n}P_\beta(\zeta,\alpha\lambda)f*E_\alpha(z\zeta^\alpha;\lambda,\beta)\,d\zeta.$$

Barnes [1, p. 291] has given the following asymptotic expansion,

$$E_\alpha(z;\lambda,\beta)\sim\frac{\alpha^{\beta-1}}{(z)^{\beta/\alpha}}e^{z^{1/\alpha}}$$

in the region $|\arg(z)|<\alpha\pi/2$, with the function tending to zero elsewhere. By choosing $f(z)=1/(1-z)$, we see as earlier that the integral formula will converge when z satisfies

$$\mathrm{Re}\,(1-z^{1/\alpha})<0.$$

In polar coordinates, this corresponds to

$$r\cos^\alpha(\theta/\alpha)<1,$$

and for $\alpha<1$, this is a hyperbola-shaped region folding around the principal star cut from $+1$ to $+\infty$ having asymptotes $\theta=\pm\alpha\pi/2$. In fact, for any given z near this edge of the principal star (i.e. r large, θ small), there will be a small value of α which makes θ/α close enough to $\pm\pi/2$ so that the above condition is satisfied and the continuation formula converges. Similar results hold for functions with other barrier points.

In conclusion, we note that in another paper to appear shortly [5] the Hadamard product on which the continuations here are based is used as a Plancherel-type formula to generate a transform calculus. The transform is used to study two dimensional, irregularly shaped semi-infinite stress systems. The results presented here are used to compute transform inversions in the continued regions by using asymptotic representations. This is similar to the approach used here to establish convergence of the various continuations.

REFERENCES

[1] E. W. BARNES, *The asymptotic expansion of integral functions defined by Taylor's series*, Philos. Trans. Roy. Soc. London Ser. A., A206 (1906), pp. 249–297.

[2] H. BATEMAN, *Higher Transcendental Functions*, vol. III, McGraw-Hill, New York, 1953.

[3] E. BOREL, *Leçons sur les Series Divergentes*, 2nd ed., Gauthier-Villars, Paris, 1928.

[4] B. F. CHAMBERS, *Hadamard's convolution integral*, Unpublished Ph.D. thesis, Univ. of Alabama, Tuscaloosa, 1969.

[5] ———, *A synthesis of transform theories with stress system applications*, to appear.

[6] J. S. HADAMARD AND M. MANDELBRAJT, *La Series de Taylor et son Prolongement Analytique*, Scienctia 41, 2nd ed., vol. I, Gauthier-Villars, Paris, 1926.

[7] E. HILLE, *Analytic Function Theory*, vol. II, Ginn, Boston, 1962.

[8] W. MAGNUS AND F. OBERHETTINGER, *Formulas and Theorems for the Special Functions of Mathematical Physics*, Chelsea, New York, 1949.

[9] G. SANSONE AND J. GERRETSEN, *Lectures on the Theory of Functions of a Complex Variable*, vol. I, P. Noordhoff, Groningen, the Netherlands, 1960.

SIAM J. MATH. ANAL.
Vol. 9, No. 6, December 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0010$01.00/0

ASYMPTOTIC BEHAVIOR FOR SEMILINEAR DIFFERENTIAL EQUATIONS IN BANACH SPACES*

R. H. MARTIN, JR.†

Abstract. Suppose that X is a Banach space and consider the integral equation

$$u(t)=T(t,a)(z-\alpha(a))+\alpha(t)+\int_a^t T(t,r)B(r,u(r))\,dr,\qquad t\geqq a\geqq 0, \tag{1}$$

where $\{T(t,s): t\geqq s\geqq 0\}$ is a linear evolution system on X, $\alpha:[0,\infty)\to X$ is continuous, $E\subset X$, and $B:[0,\infty)\times E\to X$ is continuous. In this paper we develop methods for studying the behavior as $t\to\infty$ of solutions u to (1). These results are based on Lyapunov-like methods and the proofs use standard techniques. The abstract theorems are presented in the first section and some examples indicating the applicability of these ideas are indicated in the second section. In particular, these methods are used to study the behavior of solutions to systems to semilinear parabolic equations.

1. Abstract methods. Let X be a Banach space over the real or complex field and let $|\cdot|$ denote the norm on X. In this section, the following notations and assumptions are used.

(C1) $T=\{T(t,s): t\geqq s\geqq 0\}$ is a family of bounded linear maps from X into X such that $T(t,t)=I$, $T(t,s)T(s,r)=T(t,r)$ for $t\geqq s\geqq r\geqq 0$, and the map $(t,s)\to T(t,s)x$ is continuous from $\{(t,s): t\geqq s\geqq 0\}$ into X for each $x\in X$.

(C2) $E\subset X$ and $B:[0,\infty)\times E\to X$ is continuous and bounded on bounded subsets of $[0,\infty)\times \mathrm{E}$.

(C3) α is a continuous function from $[0,\infty)$ into X.

Define the family $S_\alpha=\{S_\alpha(t,s): t\geqq s\geqq 0\}$ of bounded affine mappings on X by

$$S_\alpha(t,s)x=T(t,s)(x-\alpha(s))+\alpha(t)\quad \text{for } t\geqq s\geqq 0 \quad\text{and}\quad x\in X. \tag{1.1}$$

Now consider the semilinear equation

$$u(t)=S_\alpha(t,a)z+\int_a^t T(t,r)B(r,u(r))\,dr\quad \text{for } t\geqq a\geqq 0, \tag{1.2}$$

where $z\in E$. In this section it is assumed that

(C4) $a\in[0,\infty)$, $z\in E$, and $u:[a,\infty)\to E$ is continuous and satisfies (1.2) for all $t\geqq a$.

Conditions on T and B insuring that (C4) is satisfied may be found in [5] and [11], for example.

Our method for studying the asymptotic behavior of the solution u to (1.2) is based on standard Lyapunov-like techniques; so it is also assumed that

(C5) $V:[0,\infty)\times X\to[0,\infty)$ is continuous and for each $R>0$ there is a continuous, nondecreasing function $L_R:[0,\infty)\to[0,\infty)$ such that $|V[t,x]-V[t,y]|\leqq L_R(t)|x-y|$ for all $t\geqq 0$ and $x,y\in X$ with $|x|,|y|\leqq R$.

As a comparison equation, it is supposed that $\mu:[0,\infty)\times\mathbb{R}\to\mathbb{R}$ is continuous and that

(C6) for each $\rho_0\geqq 0$ the maximal solution $\rho(\cdot,\rho_0)$ to the initial value problem $y'=\mu(t,y)$, $y(a)=\rho_0$, $t\geqq a$, exists on $[a,\infty)$.

---

* Received by the editors January 13, 1977, and in revised form July 1, 1977.

† Department of Mathematics, North Carolina State University, Raleigh, North Carolina 27607. This work was supported in part by the U.S. Army Research Office, Research Triangle Park, North Carolina.

We have the following basic result for the growth of $t \to V[t, u(t)]$ as $t \to \infty$.

THEOREM 1. *In addition to conditions* (C1)–(C6) *suppose that*

$$\liminf_{h\to 0^+} -h^{-1}(V[t-h, u(t-h)] - V[t, S_\alpha(t, t-h)u(t-h) + hB(t-h, u(t-h))]) \tag{1.3}$$
$$\leqq \mu(t, V[t, u(t)])$$

for all $t > a$. *Then* $V[t, u(t)] \leqq \rho(t; V[a, z])$ *for all* $t \leqq a$.

Proof. Note first that if $t, h, t-h > 0$ then

$$u(t) = S_\alpha(t, t-h)u(t-h) + \int_{t-h}^{t} T(t, r)B(r, u(r))\, dr. \tag{1.4}$$

Using the continuity of T, B and u it follows that

$$\int_{t-h}^{t} T(t, r)B(r, u(r))\, dr = hB(t-h, u(t-h)) + o_1(h)$$

where $|h^{-1}o_1(h)| \to 0$ as $h \to 0^+$. Therefore, by (C5),

$$V[t, u(t)] = V[t, S_\alpha(t, t-h)u(t-h) + hB(t-h, u(t-h))] + o_2(h),$$

where $|h^{-1}o_2(h)| \to 0$ as $h \to 0^+$. Consequently, if $p(s) = V[s, u(s)]$ for all $s \geqq 0$, then

$$\begin{aligned} D_- p(t) &\equiv \liminf_{h\to 0^+} -h^{-1}(p(t-h) - p(t)) \\ &= \liminf_{h\to 0^+} -h^{-1}(V[t-h, u(t-h)] - V[t, u(t)]) \\ &\leqq \mu(t, V[t, u(t)]) = \mu(t, p(t)) \end{aligned}$$

for all $t > a$ by (1.3). Solving this differential inequality (see, e.g., [8]), we see that $p(t) \leqq \rho(t, p(a))$ for all $t \geqq a$ and the proof of this theorem is complete.

Sometimes it is more convenient to place conditions on S_α and B separately to estimate the behavior of $t \to V[t, u(t)]$ as $t \to \infty$. For this purpose define the directional Dini derivative

$$D_- V[t, x]y \equiv \liminf_{h\to 0^+} -h^{-1}(V[t, x-hy] - V[t, x]) \tag{1.5}$$

for all $x, y \in X$ and $t \geqq 0$. As a consequence of Theorem 1 we have

COROLLARY 1. *In addition to* (C1)–(C6), *suppose that* $\delta: [0, \infty) \to \mathbb{R}$ *is continuous, that*

$$V[t, S_\alpha(t, s)x] \leqq V[s, x] \exp\left(\int_s^t \delta(r)\, dr\right) \quad \text{for } t \geqq s \geqq a \quad \text{and } x \in X, \tag{1.6}$$

and that

$$D_- V[t, u(t)]B(t, u(t)) \leqq \mu(t, V[t, u(t)]) - \delta(t)V[t, u(t)] \tag{1.7}$$

for all $t > a$. *Then* $V[t, u(t)] \leqq \rho(t; V[a, z])$ *for all* $t \geqq a$.

Proof. Using (1.4) in the proof of Theorem 1 we obtain that

$$S_\alpha(t, t-h)u(t-h) = u(t) - hB(t, u(t)) + o_1(h)$$

where $h^{-1}|o_1(h)| \to 0$ as $h \to 0^+$. Therefore, by (1.6),

$$V[t-h, u(t-h)] \geqq V[t, S_\alpha(t, t-h)u(t-h)] \exp\left(-\int_{t-h}^{t} \delta(r)\, dr\right)$$
$$= V[t, u(t)-hB(t, u(t))] \exp\left(-\int_{t-h}^{t} \delta(r)\, dr\right) + o_2(h)$$

where $h^{-1}|o_2(h)| \to 0$ as $h \to 0^+$. Consequently, since

$$u(t) = S_\alpha(t, t-h)u(t-h) + hB(t-h, u(t-h)) + o_3(h),$$

we see that

$$-h^{-1}(V[t-h, u(t-h)] - V[t, S_\alpha(t, t-h)u(t-h) + hB(t-h, u(t-h))])$$
$$\leqq -h^{-1}\left(V[t, u(t)-hB(t, u(t))] \exp\left(-\int_{t-h}^{t} \delta(r)\, dr\right) - V[t, u(t)]\right) + h^{-1}o_4(h)$$
$$\leqq -h^{-1}(V[t, u(t)-hB(t, u(t))] - V[t, u(t)]) + h^{-1}o_4(h)$$
$$+ V[t, u(t)-hB(t, u(t))]\left[1-\exp\left(\int_{t-h}^{t} \delta(r)\, dr\right)\right]/h$$
$$\to D_-V[t, u(t)]B(t, u(t)) + V[t, u(t)]\delta(t)$$

as $h \to 0^+$. The corollary now follows from Theorem 1 and (1.7).

Remark 1. Note that (1.3) in Theorem 1 is certainly satisfied if

$$\liminf_{h \to 0^+} -h^{-1}(V[t-h, x] - V[t, S_\alpha(t, t-h)x + hB(t-h, x)]) \leqq \mu(t, V[t, x])$$

for all $t > a$ and $x \in E$. Similarly, (1.7) in Corollary 1 holds if

$$D_-V[t, x]B(t, x) \leqq \mu(t, V[t, x]) - \delta(t)V[t, x]$$

for all $t > a$ and $x \in E$.

Sometimes it is convenient to study the manner in which small changes in the initial conditions affect the asymptotic behavior of solutions to (1.2). So, in addition to (C1)–(C6), suppose that $w \in E$ and that $v: [a, \infty) \to E$ is continuous and

$$v(t) = S_\alpha(t, a)w + \int_a^t T(t, r)B(r, v(r))\, dr \quad \text{for all } t \geqq a;$$

that is, v is a solution to (1.2) on $[a, \infty)$ with z replaced by w.

THEOREM 2. *Suppose that the conditions enumerated in the preceding paragraph are fulfilled and also that*

$$\liminf_{h \to 0^+} -h^{-1}\{V[t-h, u(t-h)-v(t-h)] - V[t, T(t, t-h)(u(t-h)-v(t-h)) + h(B(t-h, u(t-h)) - B(t-h, v(t-h)))]\} \leqq \mu(t, V[t, u(t)-v(t)]) \tag{1.8}$$

for all $t > a$. Then $V[t, u(t)-v(t)] \leqq \rho(t; V[a, z-w])$ for all $t \geqq a$.

Indication of Proof. Define $p(t)=V[t, u(t)-v(t)]$ for all $t\geqq a$. For the proof of this theorem it suffices to show that

$$D_-p(t)\equiv\liminf_{h\to 0^+} -h^{-1}(p(t-h)-p(t))\leqq\mu(t, p(t))\quad\text{for all } t>a.$$

As in the proof of Theorem 1 we have that

$$u(t)=S_\alpha(t, t-h)u(t-h)+hB(t-h, u(t-h))+o_1(h),$$
$$v(t)=S_\alpha(t, t-h)v(t-h)+hB(t-h, u(t-h))+o_2(h),$$

and hence that

$$u(t)-v(t)=T(t, t-h)(u(t-h)-v(t-h))$$
$$+h(B(t-h, u(t-h))-B(t-h, v(t-h)))+o_3(h).$$

Therefore,

$$p(t)=V[t, T(t, t-h)(u(t-h)-v(t-h))$$
$$+h(B(t-h, u(t-h))-B(t-h, v(t-h)))]+o_4(h),$$

and it follows from (1.8) that $D_-p(t)\leqq\mu(t, p(t))$ for $t>a$, and the proof indication of Theorem 2 is complete.

COROLLARY 2. *Suppose that the conditions enumerated in the paragraph preceding Theorem 2 are fulfilled and also that δ: $[0,\infty)\to\mathbb{R}$ is a continuous function such that*

(1.9) $$V[t, T(t, s)x]\leqq V[s, x]\exp\left(\int_s^t \delta(r)\,dr\right)\quad \text{for } t\geqq s\geqq a\quad\text{and } x\in X$$

and

(1.10) $$D_-V[t, u(t)-v(t)](B(t, u(t))-B(t, v(t)))\leqq\mu(t, V[t, u(t)-v(t)])-\delta(t)V[t, u(t)-v(t)]$$

for all $t>a$. Then $V[t, u(t)-v(t)]\leqq\rho(t; V[a, z-w])$ for all $t\geqq a$.

This corollary follows from Theorem 2 in a manner analogous to the way Corollary 1 follows from Theorem 1.

Remark 2. Note that (1.8) in Theorem 2 is satisfied if

$$\liminf_{h\to 0^+} -h^{-1}\{V[t-h, x-y]-V[t, T(t, t-h)(x-y)+h(B(t-h, x)-B(t-h, y))]\}$$
$$\leqq\mu(t, V[t, x-y])$$

for all $t>a$ and $x, y\in E$. Also, (1.10) in Corollary 2 is satisfied if

$$D_-V[t, x-y](B(t, x)-B(t, y))\leqq\mu(t, V[t, x-y])-\delta(t)V[t, x-y]$$

for all $t>a$ and $x, y\in E$.

As a final abstract result, we indicate the manner in which small changes in α and B can affect the behavior of solutions to (1.2). For this result, however, we need several additional assumptions. In addition to (C1)–(C5) suppose that

(P1) β: $[0,\infty)\to X$ and k: $[0,\infty)\to X$ are continuous with

$$\lim_{t\to\infty}|\beta(t)-\alpha(t)|=\lim_{t\to\infty}|k(t)|=0;$$

(P2) $S_\beta(t, s)x \equiv T(t, s)(x-\beta(s))+\beta(t)$ for all $t \geqq s \geqq 0$ and $x \in X$, and $v: [a, \infty) \to E$ is continuous and satisfies

$$v(t)=S_\beta(t, a)w+\int_a^t T(t, r)[B(r, v(r))+k(r)]\, dr$$

for all $t \geqq a$, where $w \in E$;

(P3) there exists continuous functions γ_1, γ_2: $[0, \infty) \to X$ such that $\alpha(t)-\beta(t)=\gamma_1(t)-\gamma_2(t)$ for all $t \geqq a$, $u(t)-\gamma_1(t)$, $v(t)-\gamma_2(t) \in E$ for all $t \geqq a$, and

$$\lim_{t\to\infty} |B(t, u(t))-B(t, u(t)-\gamma(t))|$$

$$=\lim_{t\to\infty} |B(t, v(t))-B(t, v(t)-\gamma_2(t))|=0;$$

(P4) if L_R: $[0, \infty) \to [0, \infty)$ is as in (C5) then for each $R>0$ there is an $M_R>0$ such that $L_R(t) \leqq M_R$ for all $t \geqq 0$. Also, either $\sup\{M_R: R>0\}<\infty$ or $\sup\{|u(t)-v(t)|: t \geqq a\}<\infty$.

THEOREM 3. *In addition to* (C1)–(C5) *and* (P1)–(P4), *suppose that* δ, ν: $[a, \infty) \to \mathbb{R}$ *are continuous and that*

$$V[t, T(t, s)x] \leqq V[s, x] \exp\left(\int_s^t \delta(r)\, dr\right) \tag{1.11}$$

for all $t \geqq s \geqq a$ *and* $x \in X$, *that*

$$D_- V[t, x-y](B(t, x)-B(t, y)) \leqq (\nu(t)-\delta(t))V[t, x-y] \tag{1.12}$$

for all $t>a$ *and* $x, y \in E$, *and that there are numbers* N, $\nu_0>0$ *such that*

$$\exp\left(\int_s^t \nu(r)\, dr\right) \leqq N \exp(-\nu_0(t-s)) \tag{1.13}$$

for all $t \geqq s \geqq a$. *Then* $\lim_{t\to\infty} V[t, u(t)-v(t)]=0$.

For the proof of Theorem 3 we use the following lemma.

LEMMA 1. *Let the suppositions of Theorem* 3 *be fulfilled and define* $\bar{u}(t)=u(t)-\gamma_1(t)$, $\bar{v}(t)=v(t)-\gamma_2(t)$, *and* $p(t)=V[t, \bar{u}(t)-\bar{v}(t)]$ *for all* $t \geqq a$. *If*

$$D_-p(t) \equiv \liminf_{h\to 0^+} -h^{-1}[p(t-h)-p(t)]$$

for all $t>a$ *then*

$$D_-p(t) \leqq \delta(t)p(t)+D_-V[t, \bar{u}(t)-\bar{v}(t)](B(t, u(t))-B(t, v(t))-k(t))$$

for all $t>a$.

Proof. Suppose $t, h, t-h>0$. From (1.4) in the proof of Theorem 1,

$$S_\alpha(t, t-h)u(t-h)=u(t)-hB(t, u(t))+o_1(h)$$

and, similarly,

$$S_\beta(t, t-h)v(t-h)=v(t)-h[B(t, v(t))+k(t)]+o_2(h).$$

Therefore,

$$\begin{aligned} T(t, t-h)(\bar{u}(t-h)-\bar{v}(t-h)) &= S_\alpha(t, t-h)u(t-h)-\gamma_1(t)-S_\beta(t, t-h)v(t-h)+\gamma_2(t) \\ &= \bar{u}(t)-\bar{v}(t)-h\{B(t, u(t))-B(t, v(t))-k(t)\}+o_3(h) \end{aligned}$$

and it follows from (1.11) and (C5) that

$$V[t-h, \bar{u}(t-h)-\bar{v}(t-h)] \geqq V[t, T(t, t-h)(\bar{u}(t-h)-\bar{v}(t-h))] \exp\left(-\int_{t-h}^{t} \delta(r)\, dr\right)$$
$$= V[t, \bar{u}(t)-\bar{v}(t)-h\{B(t, u(t))$$
$$-B(t, v(t))-k(t)\}] \exp\left(-\int_{t-h}^{t} \delta(r)\, dr\right)+o_4(h).$$

Using this inequality along with the definition of D_-V (see (1.5)) it is easy to check that

$$D_-p(t)=\liminf_{h\to 0^+} -h^{-1}(V[t-h, \bar{u}(t-h)-\bar{v}(t-h)]-V[t, \bar{u}(t)-\bar{v}(t)])$$
$$\leqq V[t, \bar{u}(t)-\bar{v}(t)] \lim_{h\to 0^+} -h^{-1}\left\{\exp\left(-\int_{t-h}^{t} \delta(r)\, dr\right)-1\right\}$$
$$+\liminf_{h\to 0^+} -h^{-1}(V[t, \bar{u}(t)-\bar{v}(t)-h\{B(t, u(t))-B(t, v(t))-k(t)\}]$$
$$-V[t, \bar{u}(t)-\bar{v}(t)])$$
$$=V[t, \bar{u}(t)-\bar{v}(t)]\delta(t)+D_-V[t, \bar{u}(t)-\bar{v}(t)](B(t, u(t))-B(t, v(t))-k(t))$$

and the assertion of this lemma follows.

Proof of Theorem 3. It follows from the definition of D_-V and (C5) that

$$|D_-V[t, x]y-D_-V[t, x]z| \leqq L_R(t)|y-z|$$

for all $t>0$ and $x, y, z \in X$ with $|x|<R$. Consequently, (P4) implies that there is a number $M>0$ such that

$$|D_-V[t, \bar{u}(t)-\bar{v}(t)](B(t, \bar{u}(t))-B(t, \bar{v}(t)))-D_-V[t, \bar{u}(t)-\bar{v}(t)]$$
$$\cdot (B(t, u(t))-B(t, v(t))-k(t))|$$
$$\leqq M\{|B(t, \bar{u}(t))-B(t, u(t))|+|B(t, \bar{v}(t))-B(t, v(t))|+|k(t)|\}$$

for all $t>a$. Letting $\varepsilon(t)$ denote the term on the right side of this inequality, we see from Lemma 1 that

$$D_-p(t) \leqq \delta(t)p(t)+D_-V[t, \bar{u}(t)-\bar{v}(t)](B(t, \bar{u}(t))-B(t, \bar{v}(t)))+\varepsilon(t)$$
$$\leqq \nu(t)p(t)+\varepsilon(t),$$

where the last inequality follows from (1.12) and the fact that $\bar{u}(t)$, $\bar{v}(t) \in E$ for all $t>a$. Solving this differential inequality we see it follows from (1.13) that

$$V[t, \bar{u}(t)-\bar{v}(t)] \leqq NV[a, \bar{u}(a)-\bar{v}(a)]e^{-\nu_0(t-a)}+\int_a^t N\varepsilon(s)e^{-\nu_0(t-s)}\, ds$$

for all $t \geqq a$. Since $\varepsilon(s)\to 0$ as $s\to\infty$ one may easily show that $V[t, \bar{u}(t)-\bar{v}(t)]\to 0$ as $t\to\infty$. From (C5) and (P3) we have that

$$|V[t, u(t)-v(t)]-V[t, \bar{u}(t)-\bar{v}(t)]| \leqq L_R(t)|\beta(t)-\alpha(t)|$$

whenever $|u(t)-v(t)|$, $|u(t)-v(t)-\alpha(t)+\beta(t)|<R$. Using (P1) and (P4) one sees that $V[t, u(t)-v(t)]\to 0$ as $t\to\infty$ and the proof of Theorem 3 is complete.

2. Examples. In this section we indicate the applicability of the abstract methods in § 1 to semilinear parabolic systems. For our first example it is supposed that Ω is a smooth bounded domain in $\mathbb{R}^n$ and that Δ is the Laplacian operator on Ω. Also, m is a positive integer and $\mathbb{R}_+^m$ is the positive cone in $\mathbb{R}^m$: $\mathbb{R}_+^m=\{\xi=(\xi_i)_1^m \in \mathbb{R}^m: \xi_i \geqq 0$ for $i=1, \cdots, m\}$. Now suppose that a_i and R_i are positive numbers for each $i=1, \cdots, m$ and define the class $A_+(a, R)$ of all continuous functions $\alpha=(\alpha_i)_1^m$ from $[0, \infty)\times \bar{\Omega}$ into $\mathbb{R}^m$ by

$$\begin{aligned} &\alpha \in \mathscr{A}_+(a, R) \quad \text{only in case } 0 \leqq \alpha_i(t, \tau) \leqq R_i \quad \text{and} \\ &\frac{\partial}{\partial t}\alpha_i(t, \tau)=a_i\Delta\alpha_i(t, \tau) \quad \text{for all } (t, \tau) \in (0, \infty)\times\Omega, \quad i=1, \cdots, m. \end{aligned} \tag{2.1}$$

Note that each $\alpha \in \mathscr{A}_+(a, R)$ is uniquely determined by its values on $\{0\}\times\Omega$ and $[0, \infty)\times\partial\Omega$.

Suppose that $f=(f_i)_1^m$ is a continuous function from $\mathbb{R}_+^m$ into $\mathbb{R}^m$. We study the asymptotic behavior of nonnegative, uniformly bounded solutions to the semilinear parabolic system

$$\begin{aligned} &\frac{\partial}{\partial t}u_i(t, \tau)=a_i\Delta u_i(t, \tau)+f_i(u_1(t, \tau), \cdots, u_m(t, \tau)) \\ &u_i(0, \tau)=\chi_i(\tau) \quad \text{and} \quad u_i(t, \sigma)=\alpha_i(t, \sigma), \\ &\text{for } t>0, \quad \tau \in \Omega, \quad \sigma \in \partial\Omega, \quad i=1, \cdots, m \end{aligned} \tag{2.2}$$

where we make the following basic assumptions

(A1) χ_i is measurable from Ω into $[0, R_i]$ for each $i=1, \cdots, m$.

(A2) $\alpha=(\alpha_i)_1^m \in \mathscr{A}_+(a, R)$.

(A3) f has the property that if $j \in \{1, \cdots, m\}$ and $\xi=(\xi_i)_1^m \in \mathbb{R}^m$ with $0 \leqq \xi_i \leqq R_i$ for each $i=1, \cdots, m$, then $\xi_j=0$ implies $f_j(\xi)\geqq 0$ and $\xi_j=R_j$ implies $f_j(\xi)\leqq 0$.

Define $|\xi|_p=(\sum_{i=1}^m |\xi_i|^p)^{1/p}$ for all $\xi \in \mathbb{R}^m$ and $p \in [1, \infty)$, and denote by $X_p(=\mathscr{L}^p(\Omega; \mathbb{R}^m))$ the space of all measurable functions $\phi=(\phi_i)_1^m: \Omega \to \mathbb{R}^m$ such that

$$\|\phi\|_p \equiv \left[\int_\Omega |\phi(\tau)|_p^p \, d\tau\right]^{1/p} < \infty.$$

For each subset Λ of $\mathbb{R}^m$ let

$$K_p(\Lambda)=\{\phi \in X_p: \phi(\tau) \in A \text{ for almost all } \tau \in \Omega\}.$$

Now define the operators L_p and B_p on X_p by

$$\begin{aligned} &L_p\phi=(a_i\Delta\phi_i)_1^m \quad \text{for all } \phi=(\phi_i)_1^m \in D(L_p) \quad \text{where} \\ &D(L_p)=\{\phi \in X_p: \phi_i \in H^{2,p}(\Omega)\cap H_0^{1,p}(\Omega) \text{ for } i=1, \cdots, m\} \end{aligned} \tag{2.3}$$

where $H^{2,p}(\Omega)$ and $H_0^{1,p}(\Omega)$ are the usual Sobolev spaces (see, e.g., [4, p. 14 and p. 33]) and

$$\begin{aligned} &[B_p\phi](\tau)=f(\phi(\tau)) \quad \text{for all } \tau \in \Omega \quad \text{and} \quad \phi \in D(B_p) \quad \text{where} \\ &D(B_p)=\{\phi \in K_p(\mathbb{R}_+^m): \tau \to f(\phi(\tau)) \text{ is in } X_p\}. \end{aligned} \tag{2.4}$$

It is well known that L_p is the generator of a compact, analytic semigroup $T_p=\{T_p(t): t \geqq 0\}$ of bounded linear operators on X_p (see, e.g., Friedman [4] or Pazy [12]).

Moreover,

(2.5) $$\|T_p(t)\phi\|_p \leqq \|\phi\|_p e^{\omega_p t} \quad \text{for all } t \geqq 0 \quad \text{and } \phi \in X_p,$$

where $\omega_1 = 0$ and $\omega_p < 0$ if $p > 1$. For each $\chi \in D(B_p)$ and $\alpha \in \mathcal{A}_+(a, R)$ consider the integral equation

(2.6) $$u(t) = T_p(t)(\chi - \alpha(0, \cdot)) + \alpha(t, \cdot) + \int_0^t T_p(t-r)B_p u(r)\, dr$$

for $t \geqq 0$. Solutions $u: [0, \infty) \to X_p$ to (2.6) are called mild X_p-solutions to (2.2). We have the following existence result for solutions to (2.6):

LEMMA 2. *Suppose that* (A1)–(A3) *are satisfied,* $1 \leqq p < \infty$, T_p *and* B_p *are as above, and* $\Lambda = \prod_{i=1}^m [0, R_i] \subset \mathbb{R}_+^m$. *Then for each* $\chi \in K_p(\Lambda)$ *and* $\alpha \in \mathcal{A}_+(a, R)$ *there is a solution* $u = u_{\alpha,\chi}$ *to* (2.6) *on* $[0, \infty)$ *such that* $u_{\alpha,\chi}(t) \in K_p(\Lambda)$ *for all* $t \geqq 0$.

This lemma follows from Proposition 5.1 of [11] (see also Remarks 3.1 and 5.2 in [11]). Our main result on the asymptotic behavior of mild X_p-solutions to (2.2) is

PROPOSITION 1. *Suppose that* (A1) *and* (A3) *are satisfied,* $1 \leqq p < \infty$, and $\Lambda = \prod_{i=1}^m [0, R_i]$. *Suppose further that*

(2.7) $$|\xi - \eta - h[f(\xi) - f(\eta)]|_p \geqq |\xi - \eta|_p \quad \text{for all } \xi, \eta \in \mathbb{R}_+^m \quad \text{and} \quad h > 0$$

and that $\alpha, \beta \in \mathcal{A}_+(a, R)$ *with*

(2.8) $$\lim_{t\to\infty} \|\alpha(t, \cdot) - \beta(t, \cdot)\|_p = 0.$$

Then $\lim_{t\to\infty} \|u_{\alpha,\chi}(t) - u_{\beta,\psi}(t)\|_p = 0$ *for each* $\chi, \psi \in K_p(\Lambda)$.

Proof. Suppose first that $1 < p < \infty$ and define $V[\phi] = \|\phi\|_p$ for all $\phi \in X_p$. From (2.5) it follows that (1.11) in Theorem 3 is satisfied with $\delta(r) \equiv \omega_p$ for all $r \geqq 0$. Also, (2.7) implies that

$$|\phi(\tau) - \bar\phi(\tau) - h[B_p\phi(\tau) - B_p\bar\phi(\tau)]|_p^p \geqq |\phi(\tau) - \bar\phi(\tau)|_p^p$$

for all $h > 0$, $\tau \in \Omega$, and integrating each side of this inequality over Ω shows that

$$V[\phi - \bar\phi - h[B_p\phi - B_p\bar\phi]] \geqq V[\phi - \bar\phi] \quad \text{for all } \phi, \quad \bar\phi \in D(B_p), \quad h > 0.$$

Thus (1.12) is satisfied with $\nu(t) \equiv \omega_p$. Since $\omega_p < 0$, (1.13) is satisfied as well. Setting $\xi_+ = ([\xi_i + |\xi_i|]/2)_1^m$ for each $\xi = (\xi_i)_1^m \in \mathbb{R}^m$ and defining

$$\gamma_1(t, \tau) = -[\beta(t, \tau) - \alpha(t, \tau)]_+ \quad \text{and} \quad \gamma_2(t, \tau) = -[\alpha(t, \tau) - \beta(t, \tau)]_+$$

for all $(t, \tau) \in [0, \infty) \times \Omega$, one sees that $\alpha - \beta = \gamma_1 - \gamma_2$, $u_{\alpha,\chi}(t) - \gamma_1(t, \cdot) \in D(B_p)$ and $u_{\beta,\psi}(t) - \gamma_2(t, \cdot) \in D(B_p)$ for all $t \geqq 0$. Also,

$$\lim_{t\to\infty} \|\gamma_1(t, \cdot)\|_p = \lim_{t\to\infty} \|\gamma_2(t, \cdot)\|_p = 0.$$

Choosing R' sufficiently large so that

$$u_{\alpha,\chi}(t) - \gamma_1(t, \cdot),\ u_{\beta,\psi}(t) - \gamma_2(t, \cdot) \in K_p(\Lambda')$$

where $\Lambda' = \{\xi \in \mathbb{R}_+^m : |\xi|_p \leqq R'\}$, one may easily check that each of the conditions of Theorem 3 are satisfied with $E = K_p(\Lambda')$. Thus $\|u_{\alpha,\chi}(t) - u_{\beta,\psi}(t)\|_p \to 0$ on $t \to \infty$ by Theorem 3 and the proof of this proposition is complete for $p > 1$.

For $p = 1$ we use the techniques of Kahane in [6, pp. 353–354]. Let λ_1 denote the first eigenvalue of the problem $\Delta\phi + \lambda\phi = 0$ in Ω and $\phi = 0$ on $\partial\Omega$. Then $\lambda_1 > 0$ and there is a corresponding eigenfunction Φ_1 that has positive values on Ω (see [7, p. 259]). Now define

$$V[\phi]=\int_\Omega \Phi_1(\tau)|\phi(\tau)|_1\,d\tau \quad \text{for all } \phi\in X_1.$$

Since Φ_1 is continuous and $\Phi_1(\tau)>0$ for $\tau\in\Omega$ and $u_{\alpha,\chi}$ and $u_{\beta,\psi}$ are uniformly bounded on $[0,\infty)\times\Omega$, it suffices to show that $V[u_{\alpha,\chi}(t)-u_{\beta,\psi}(t)]\to 0$ as $t\to\infty$. From (2.7) it follows that

$$\Phi_1(\tau)|\phi(\tau)-\bar\phi(\tau)-h[B_1\phi(\tau)-B_1\bar\phi(\tau)]|_1\geqq \Phi_1(\tau)|\phi(\tau)-\bar\phi(\tau)|_1$$

for all $\tau\in\Omega$ and $h>0$, and hence it is easy to see that

$$V[\phi-\bar\phi-h(B_1\phi-B_1\bar\phi)]\geqq V[\phi-\bar\phi] \quad \text{for all } \phi,\ \bar\phi\in D(B_1),\quad h>0.$$

Therefore, $D_-V[\phi-\bar\phi](B_1\phi-B_1\bar\phi)\leqq 0$ for all $\phi,\bar\phi\in D(B_1)$, and the assertions in this case will follow as in the case when $p>1$ once it is shown that

$$V[T_1(t)\phi]\leqq V[\phi]\,e^{-\delta t} \quad \text{for all } \phi\in X_1 \quad\text{and}\quad t\geqq 0, \tag{2.9}$$

where $\delta>0$. Now suppose $\phi:\bar\Omega\to\mathbb{R}^m_+$ is continuous and define $v(t,\tau)=[T_1(t)\phi](\tau)$ for all $(t,\tau)\in[0,\infty)\times\bar\Omega$. Then $v(t,\tau)=(v_i(t,\tau))_1^m\in\mathbb{R}^m_+$ and it follows that

$$\frac{d}{dt}V[v(t,\cdot)]=\sum_{i=1}^m\int_\Omega \Phi_1(\tau)\frac{\partial}{\partial t}v_i(t,\tau)\,d\tau=\sum_{i=1}^m\int_\Omega \Phi_1(\tau)a_i\Delta v_i(t,\tau)\,d\tau.$$

Since $\Phi_1(\sigma)=v_i(t,\sigma)=0$ for $\sigma\in\partial\Omega$ and $\Delta\Phi_1+\lambda_1\Phi_1=0$ in Ω, it follows that

$$\int_\Omega \Phi_1(\tau)a_i\Delta v_i(t,\tau)\,d\tau=a_i\int_\Omega [\Delta\Phi_1(\tau)]v_i(t,\tau)\,d\tau=-\lambda_1 a_i\int_\Omega \Phi_1(\tau)v_i(t,\tau)\,d\tau,$$

and hence if $\delta=\min\{\lambda_1 a_i: i=1,\cdots,m\}$ then $\delta>0$ and

$$\frac{d}{dt}V[v(t,\cdot)]=\sum_{i=1}^m-\lambda_1 a_i\int_\Omega \Phi_1(\tau)v_i(t,\tau)\,d\tau\leqq -\delta V[v(t,\cdot)].$$

Solving this differential inequality shows that (2.9) holds for all continuous $\phi:\bar\Omega\to\mathbb{R}^m_+$ and hence for all $\phi\in K_1(\mathbb{R}^m_+)$. Defining

$$\phi_+(\tau)=([|\phi_i(\tau)|+\phi_i(\tau)]/2)_1^m \quad\text{and}\quad \phi_-(\tau)=([|\phi_i(\tau)|-\phi_i(\tau)]/2)_1^m$$

for all $\phi\in X_1$ and $\tau\in\Omega$, one sees that $\phi_+,\phi_-\in K_1(\mathbb{R}^m_+)$, $\phi=\phi_+-\phi_-$, and $V[\phi]=V[\phi_+]+V[\phi_-]$. Thus, if $\phi\in X_1$,

$$\begin{aligned}V[T_1(t)\phi]&=V[T_1(t)\phi_+-T_1(t)\phi_-]\leqq V[T_1(t)\phi_+]+V[T_1(t)\phi_-]\\&\leqq (V[\phi_+]+V[\phi_-])\,e^{-\delta t}=V[\phi]\,e^{-\delta t},\end{aligned}$$

and hence (2.9) holds for all $\phi\in X_1$ and $t\geqq 0$. This completes the proof of Proposition 1.

Our second result for mild $\mathscr{L}^p$-solutions to (2.2) uses the technique in [10], and is concerned with the existence and asymptotic stability of an equilibrium solution to (2.2) when α is time independent.

PROPOSITION 2. *Suppose that* (A1) *and* (A3) *are satisfied,* $1\leqq p<\infty$, $\Lambda=\prod_{i=1}^m[0,R_i]$, *and* $\gamma\in\mathscr{A}_+(a,R)$ *is independent of* $t\geqq 0$. *Then there is a member* ψ *of* $K_p(\Lambda)$ *such that* $u_{\gamma,\psi}(t)\equiv\psi$ *for all* $t\geqq 0$, *and hence* $\psi-\gamma\in D(L_p)$ *and*

$$a_i\Delta\psi_i+f_i(\psi)=\theta \quad \textit{a.e. on } \Omega \textit{ for } i=1,\cdots,m.$$

If f also satisfies (2.7) *in Proposition* 1 *then there is precisely one $\psi \in K_p(\Lambda)$ such that $u_{\gamma,\psi}(t) = \psi$ for all $t \geqq 0$ and, moreover,*

$$\lim_{t\to\infty} \|u_{\alpha,\chi}(t) - \psi\|_p = 0$$

whenever $\chi \in K_p(\Lambda)$ and $\alpha \in \mathscr{A}_+(a, R)$ with $\lim_{t\to\infty} \|\alpha(t,\cdot) - \psi\|_p = 0$.

Proof. Define the operator M_γ on X_p by $D(M_\gamma) = \{\phi \in X_p: \phi - \gamma \in D(L_p)\}$ and $M_\gamma\phi = (a_i \Delta \phi_i)_1^m$ for all $\phi \in D(M_\gamma)$. If $h > 0$ and $\psi \in K_p(\Lambda)$ it follows easily from the maximum principle that the solution $\phi \in D(M_\gamma)$ to the equation $\phi - hM_\gamma\phi = \psi$ is also in $K_p(\Lambda)$. Hence $(I - hM_\gamma)^{-1}: K_p(\Lambda) \to K_p(\Lambda)$ for each $h > 0$. Since (A3) implies

$$\lim_{h\to 0^+} d(\xi + hf(\xi); \Lambda)/h = 0 \quad \text{for all } \xi \in \Lambda,$$

we have from [10, Theorem 1] that there is a $\phi \in K_p(\Lambda) \cap D(M_\gamma)$ such that $M_\gamma\psi + B_p\psi = \theta$. Thus $u_{\gamma,\psi}(t) \equiv \psi$ for all $t \geqq 0$. Since the final assertion is an immediate consequence of Proposition 1, the proof of Proposition 2 is complete.

Remark 3. Under additional smoothness assumptions one can conclude that the convergence of $u_{\alpha,\chi}(t) - u_{\beta,\psi}(t)$ to zero in Proposition 1 is uniform for $\tau \in \Omega$. Suppose that the map f is Hölder continuous on each bounded subset of $\mathbb{R}_+^m$ and also that the maps $t \to \alpha(t,\cdot)$ and $t \to \beta(t,\cdot)$ are continuously differentiable from $[0,\infty)$ into X_q for each $q \geqq 1$. Now let each of the suppositions of Proposition 1 be satisfied with (2.8) replaced by

$$\lim_{t\to\infty} \max\{|\alpha(t,\sigma) - \beta(t,\sigma)|_p: \sigma \in \partial\Omega\} = 0. \tag{2.8$'$}$$

If $\|\phi\|_\infty \equiv \operatorname{ess\,sup}\{|\phi(\tau)|_p: \tau \in \Omega\}$ for each essentially bounded member ϕ of X_p, then one can conclude that $\lim_{t\to\infty} \|u_{\alpha,\chi}(t) - u_{\beta,\psi}(t)\|_\infty = 0$ for each $\chi, \psi \in K_p(\Lambda)$. We give an indication of the proof of this assertion. Note first that (2.8)$'$ implies that

$$\lim_{t\to\infty} \|\alpha(t,\cdot) - \beta(t,\cdot)\|_\infty = 0 \tag{2.8$''$}$$

(see [3, p. 157]), and hence the conclusions of Proposition 1 are fulfilled. Therefore, to establish the uniform convergence of $u_{\alpha,\chi}(t) - u_{\beta,\psi}(t)$ to zero, it suffices (by Ascoli's theorem) to show that the families $\{u_{\alpha,\chi}(t): t \geqq 0\}$ and $\{u_{\beta,\psi}(t): t \geqq 0\}$ are uniformly equicontinuous on Ω, since we already have that they are uniformly bounded. Note that $u_{\alpha,\chi}$ is a mild X_q solution to (2.2) for all $q \geqq 1$. Since f (and hence B_q) is Hölder continuous it follows that

$$\int_0^t T_q(t-r)B_q u_{\alpha,\chi}(r)\, dr \in D(L_q) \quad \text{for all } t > 0.$$

(See Pazy [13]). From this it follows that if $\frac{1}{2} < \rho < 1$ and $(-L_q)^\rho$ is the fractional power of $-L_q$ (see [4] or [12]), then

$$(-L_q)^\rho \int_0^t T_q(t-r)B_q u_{\alpha,\chi}(r)\, dr = \int_0^t (-L_q)^\rho T_q(t-r)B_q u_{\alpha,\chi}(r)\, dr$$

for all $t > 0$. Since there is a constant $C_\rho > 0$ such that

$$\|(-L_q)^\rho T_q(t-r)\phi\|_q \leqq C_\rho (t-r)^{-\rho} e^{-\omega_q(t-r)} \|\phi\|_q$$

for all $t > r \geqq 0$ and $\phi \in X_q$ (see [4, p. 160]) and since $\|B_q u_{\alpha,\chi}(r)\|_q$ is uniformly bounded for $r \geqq 0$, it follows that there is a number $M_{q,\rho} > 0$ such that

$$\|(-L_q)^\rho \int_0^t T_q(t-r)B_q u_{\alpha,\chi}(r)\,dr\|_q \leqq M_{q,\rho} \quad \text{for all } t,\, q > 1.$$

The same inequality holds with $u_{\alpha,\chi}$ replaced by $u_{\beta,\psi}$. Taking q sufficiently large shows that the families

$$\left\{\int_0^t T_q(t-r)B_q u_{\alpha,\chi}(r)\,dr \colon t \geqq 1\right\} \quad \text{and} \quad \left\{\int_0^t T_q(t-r)B_q u_{\beta,\psi}(r)\,dr \colon t \geqq 1\right\}$$

are uniformly equicontinuous on Ω. Since

$$T_q(t)\phi = T_p(t)\phi \quad \text{and} \quad B_q\phi = B_p\phi \quad \text{if } \phi \in K_p(\mathbb{R}^m_+) \quad \text{with } \|\phi\|_\infty < \infty,$$

and since

$$\|T_p(t)(\chi - \alpha(0,\cdot)) + \alpha(t,\cdot) - T_p(t)(\psi - \beta(0,\cdot)) - \beta(t,\cdot)\|_\infty \to 0$$

as $t \to \infty$ it now follows easily that $\|u_{\alpha,\chi}(t) - u_{\beta,\psi}(t)\|_\infty \to 0$ as $t \to \infty$ and the assertion is established.

Remark 4. Consider the m-dimensional differential system

$$x'(t) = f(x(t)), \qquad t \geqq 0, \qquad x(0) = \xi \in \mathbb{R}^m_+. \tag{2.10}$$

The condition (A3) is equivalent to requiring that for each $\xi \in \Lambda$ $(= \prod_{i=1}^m [0, R_i])$ equation (2.10) has a solution x with $x(t) \in \Lambda$ for $t \geqq 0$. If $a_1 = \cdots = a_m$ then Propositions 1 and 2 remain valid with Λ any compact, convex subset of $\mathbb{R}^m_+$ with nonempty interior. In this case condition (A3) is replaced by: (A3)′ for each $\xi \in \Lambda$ equation (2.10) has a solution x with $x(t) \in \Lambda$ for $t \geqq 0$. The dissipative condition (2.7) is also related to equation (2.10). Assume that for each $\xi \in \mathbb{R}^m_+$, (2.10) has a solution $x_\xi \colon [0, \infty) \to \mathbb{R}^m_+$. Then (2.7) is equivalent to requiring that $|x_\xi(t) - x_\eta(t)|_p \leqq |\xi - \eta|_p$ for all $t \geqq 0$ and $\xi, \eta \in \mathbb{R}^m_+$. If f is C^1 and $df(\xi)$ denotes the differential of f at $\xi \in \mathbb{R}^m_+$, then (2.7) is equivalent to

$$|\eta - h\, df(\xi)\eta|_p \geqq 1 \quad \text{for all } \xi \in \mathbb{R}^m_+ \text{ and } \eta \in \mathbb{R}^m \quad \text{with } |\eta|_p = 1. \tag{2.7$'$}$$

This provides a convenient means of checking when (2.7) is satisfied.

Propositions 1 and 2 (along with Remark 3) contain the results of Kahane [6]. Here the system

$$\frac{\partial}{\partial t} u_1(t, \tau) = a_1 \Delta u_1(t, \tau) - u_1(t, \tau)u_2(t, \tau),$$

$$\frac{\partial}{\partial t} u_2(t, \tau) = a_2 \Delta u_2(t, \tau) - u_1(t, \tau)u_2(t, \tau)$$

is studied with bounded, nonnegative initial and boundary conditions for u_1 and u_2 prescribed on $\{0\} \times \Omega$ and $[0, \infty) \times \partial\Omega$ (see equation (3.1) of [6]). In the preceding equation the function f has the form $f(\xi_1, \xi_2) = (-\xi_1\xi_2, -\xi_1\xi_2)$ for all $(\xi_1, \xi_2) \in \mathbb{R}^2_+$; and it is easy to see that (A3) holds for any $R_1, R_2 > 0$. Moreover, (2.7) holds for $p = 1$ (but not for $p > 1$). One should also note that Propositions 1 and 2 (with $p = 1$) apply to this equation with nonlinear terms $-u_1u_2$ replaced by $\gamma(u_1, u_2)$ where $\gamma \colon \mathbb{R}^2_+ \to \mathbb{R}$ is continuously differentiable, $\gamma(0, \xi_2) = \gamma(\xi_1, 0) = 0$, $(\partial/(\partial\xi_1))\gamma(\xi_1, \xi_2) \leqq 0$ and $(\partial/(\partial\xi_2))\gamma(\xi_1, \xi_2) \leqq 0$ for all $(\xi_1, \xi_2) \in \mathbb{R}^2_+$. As a second illustration, define the map f on $\mathbb{R}^2_+$ by $f(\xi_1, \xi_2) = (-\xi_1\xi_2, -\xi_2 + \xi_1^2/2)$. Then (A3) holds for any $R_1, R_2 > 0$ with $R_2 \geqq R_1^2/2$, and (2.7) holds for $p = 2$ (but not for $p = 1$).

The form of the preceding equation is analogous to an equation arising in chemical reactions. Following Danckwerts [2], Brian, Hurley and Hasseltine [1], and Pearson [14], we consider the system

$$
(2.11)\qquad
\begin{aligned}
&\frac{\partial}{\partial t}\mu(t,\tau)=a_1\frac{\partial^2}{\partial\tau^2}\mu(t,\tau)-b_1\mu(t,\tau)\nu(t,\tau),\\
&\frac{\partial}{\partial t}\nu(t,\tau)=a_2\frac{\partial^2}{\partial\tau^2}\nu(t,\tau)-b_2\mu(t,\tau)\nu(t,\tau), \qquad \text{for } t,\tau>0,\\
&\mu(0,\tau)=0, \qquad \nu(0,\tau)=c_2, \qquad \mu(t,0)=c_1,\\
&\frac{\partial}{\partial\tau}\nu(t,0)=0, \qquad \mu(t,\infty)=0, \qquad \nu(t,\infty)=c_2,
\end{aligned}
$$

where $a_i, b_i, c_i>0$ for $i=1,2$. In place of this equation we consider the following more general system:

$$
(2.12)\qquad
\begin{aligned}
&\frac{\partial}{\partial t}u_1(t,\tau)=a_1\frac{\partial^2}{\partial\tau^2}u_1(t,\tau)-\gamma(u_1(t,\tau),u_2(t,\tau)),\\
&\frac{\partial}{\partial t}u_2(t,\tau)=a_2\frac{\partial^2}{\partial\tau^2}u_2(t,\tau)+\gamma(u_1(t,\tau),u_2(t,\tau)),\\
&u_1(0,\tau)=\chi_1(\tau), \qquad u_2(0,\tau)=x_2(\tau), \qquad u_1(t,0)=\sigma,\\
&\frac{\partial}{\partial\tau}u_2(t,0)=0, \qquad u_1(t,\infty)=0, \qquad u_2(t,\infty)=0,
\end{aligned}
$$

where $a_1, a_2>0$, $\sigma\geqq 0$, and $\chi=(\chi_1,\chi_2)$: $[0,\infty)\to\mathbb{R}_+^2$ is measurable and essentially bounded. Also, $\gamma:\mathbb{R}_+^2\to\mathbb{R}$ is continuously differentiable and there are numbers $\rho_1,\rho_2>0$ such that

($\gamma 1$) $\rho_1\geqq\sigma$,
($\gamma 2$) if $(\xi_1,\xi_2)\in[0,\rho_1]\times[0,\rho_2]$ then $\gamma(0,\xi_2)=\gamma(\xi_1,\rho_2)=0$,
($\gamma 3$) $(\partial/(\partial\xi_1))\gamma(\xi_1,\xi_2)\geqq 0$ and $(\partial/(\partial\xi_2))\gamma(\xi_1,\xi_2)\leqq 0$ if $(\xi_1,\xi_2)\in[0,\rho_1]\times[0,\rho_2]$.

A typical function γ satisfying ($\gamma 1$)–($\gamma 3$) is $\gamma(\xi_1,\xi_2)=\xi_1^{d_1}|\rho_2-\xi_2|^{d_2}$ where $d_1, d_2\geqq 1$. Equation (2.11) is subsumed by (2.12) by taking $\chi_1(\tau)=\chi_2(\tau)=0$ for $\tau\geqq 0$, $\sigma=\rho_1=b_2c_1$, $\rho_2=b_1c_2$, $\gamma(\xi_1,\xi_2)=\xi_1(\rho_2-\xi_2)$ for all $\xi_1,\xi_2\geqq 0$, and using the change of variables

$$u_1(t,\tau)=b_2\mu(t,\tau) \quad\text{and}\quad u_2(t,\tau)=b_1c_2-b_1\nu(t,\tau)$$

for all $t,\tau\geqq 0$.

The system (2.12) is analyzed in the space $\mathcal{L}^1$ $(=\mathcal{L}^1([0,\infty):\mathbb{R}^2))$ of all measurable functions $\phi=(\phi_1,\phi_2)$: $[0,\infty)\to\mathbb{R}^2$ such that

$$\|\phi\|_1\equiv\int_0^\infty|\phi(\tau)|_1\,d\tau\left(=\int_0^\infty[|\phi_1(\tau)|+|\phi_2(\tau)|]\,d\tau\right)<\infty.$$

Define the operator L on $\mathcal{L}^1$ by

$$
(2.13)\qquad
\begin{aligned}
&L\phi=(a_1\phi_1'',a_2\phi_2'') \quad \text{for all } \phi=(\phi_1,\phi_2)\in D(L) \quad \text{where}\\
&D(L)=\{\phi\in\mathcal{L}^1:\ \phi,\phi' \text{ are loc. abs. cont.}, \phi''\in\mathcal{L}^1, \phi_1(0)=\phi_2'(0)=0\}.
\end{aligned}
$$

One should note that

$$(2.14)\qquad \lim_{\tau\to\infty}|\phi(\tau)|_1=\lim_{\tau\to\infty}|\phi'(\tau)|_1=0 \quad \text{for all } \phi\in D(L).$$

It follows that L is the generator of an analytic (C_0) semigroup $T=\{T(t): t \geqq 0\}$ of bounded linear operators on $\mathscr{L}^1$ and also that

(2.15) $$\|T(t)\phi\|_1 \leqq \|\phi\|_1 \quad \text{for all } t \geqq 0 \quad \text{and} \quad \phi \in \mathscr{L}^1.$$

Set $D=\{\phi \in \mathscr{L}^1: \phi(\tau) \in [0, \rho_1] \times [0, \rho_2] \text{ a.e. } \tau \in [0, \infty)\}$ and define the operator B from D into $\mathscr{L}^1$ by

(2.16) $$[B\phi](\tau)=(-\gamma(\phi(\tau)), \gamma(\phi(\tau))) \quad \text{for all } \tau \geqq 0 \quad \phi \in D(B) \equiv D.$$

Finally let $\alpha=(\alpha_1, \alpha_2)$: $[0, \infty) \to \mathbb{R}_+^2$ be defined by $\alpha_2(t, \tau) \equiv 0$ and α_1 is the (unique) bounded solution to

$$\frac{\partial}{\partial t}\alpha_1(t, \tau)=a_1\frac{\partial^2}{\partial \tau^2}\alpha_1(t, \tau), \qquad \alpha_1(0, \tau)=0, \qquad \alpha_1(t, 0)=\sigma, \qquad \alpha_1(t, \infty)=0$$

for $t, \tau>0$.

In particular, $\alpha_1(t, \tau)=\sigma[1-\operatorname{erf}(a_1^{-1/2}t^{-1/2}\tau/2)]$ for all $t, \tau>0$ where

$$\operatorname{erf}(s)=2\pi^{-1/2}\int_0^s e^{-r^2}\,dr \quad \text{for all } s \geqq 0.$$

In the space $\mathscr{L}^1$ consider the integral equation

(2.17) $$u_\chi(t)=T(t)(\chi-\alpha(0, \cdot\,))+\alpha(t, \cdot\,)+\int_0^t T(t-r)Bu_\chi(r)\,dr \quad \text{for } t \geqq 0.$$

As before, if u_χ is a solution to (2.17) and $(u_1(t, \tau), u_2(t, \tau)) \equiv [u_\chi(t)](\tau)$ then (u_1, u_2) is a mild $\mathscr{L}^1$-solution to (2.12). Define the partial ordering "$\geqq$" on $\mathscr{L}^1$ by $\phi \geqq \psi$ if and only if $\phi_1(\tau) \geqq \psi_1(\tau)$ and $\phi_2(\tau) \geqq \psi_2(\tau)$ for almost all $\tau \in [0, \infty)$, and define the operator A on $\mathscr{L}^1$ by $A\phi=(a_1\phi_1'', a_2\phi_2'')$ for all $\phi=(\phi_1, \phi_2) \in D(A)$ where $D(A)=\{\phi \in \mathscr{L}^1: \phi, \phi'$ are loc. abs. cont., $\phi'' \in \mathscr{L}^1, \phi_1(0)=\sigma, \phi_2'(0)=0\}$. We have the following fundamental result for the existence and behavior of solutions to (2.17) (recall that $D=\{\phi \in \mathscr{L}^1: \phi(\tau) \in [0, \rho_1] \times [0, \rho_2] \text{ a.e. } \tau \geqq 0\}$).

PROPOSITION 3. *Suppose that the conditions and notations enumerated in the preceding three paragraphs are satisfied. Then for each $\chi \in D$ there is a unique solution u_χ to* (2.17) *on $[0, \infty)$ such that $u_\chi(t) \in D$ for all $t \geqq 0$. Also, if the family $U=\{U(t): t \geqq 0\}$ of mappings from D into D is defined by $U(t)\chi=u_\chi(t)$ for all $t \geqq 0$ and $\chi \in D$, then U is a (C_0) semigroup of nonlinear operators on D and U is nonexpansive, order preserving, and differentiable*:

(i) $U(0)\chi=\chi$ *and* $U(t)U(s)\chi=U(t+s)\chi$ *for all* $\chi \in D$, $t, s \geqq 0$.

(ii) $t \to U(t)\chi$ *is continuous on* $[0, \infty)$ *for each* $\chi \in D$.

(iii) $\|U(t)\chi-U(t)\psi\|_1 \leqq \|\chi-\psi\|_1$ *for all* $t \geqq 0$, $\chi, \psi \in D$.

(iv) $U(t)\chi \geqq U(t)\psi$ *whenever* $t \geqq 0$ *and* $\chi, \psi \in D$ *with* $\chi \geqq \psi$.

(v) $U(t)\chi \in D(A)$ *and* $(d/(dt))U(t)\chi=AU(t)\chi+BU(t)\chi$ *for all* $t>0$ *and* $\chi \in D$.

Moreover, if $\sigma>0$ then $\lim_{t\to\infty}\|U(t)\chi\|_1=\infty$ for each $\chi \in D$.

Indication of Proof. Since $\rho_1 \geqq \sigma$ the maximum principle implies that $T(t-s)(\chi-\alpha(s, \cdot\,))+\alpha(t, \cdot\,) \in D$ whenever $\chi \in D$ and $t \geqq s \geqq 0$. Also, (γ_2) and $(\gamma 3)$ imply that

$$\lim_{h\to 0^+} d(\xi+h(-\gamma(\xi), \gamma(\xi)); [0, \rho_1] \times [0, \rho_2])/h=0$$

for all $\xi \in [0, \rho_1] \times [0, \rho_2]$, and since B is Lipschitz continuous we have the existence and uniqueness of solutions to (2.17) on $[0, \infty)$ by [11, Thm. 3.2] and [9, Prop. 5]. This shows that U satisfies (i) and (ii) and the fact that T is analytic may be used to show that U also satisfies (v) (see Pazy [13]). Note that the Jacobian matrix $df(\xi)$ of the map $f(\xi)\equiv(-\gamma(\xi), \gamma(\xi))$ has the form

$$df(\xi)=\begin{bmatrix} -\dfrac{\partial}{\partial\xi_1}\gamma(\xi) & -\dfrac{\partial}{\partial\xi_2}\gamma(\xi) \\ \dfrac{\partial}{\partial\xi_1}\gamma(\xi) & \dfrac{\partial}{\partial\xi_2}\gamma(\xi) \end{bmatrix} \quad \text{for all } \xi\in[0,\rho_1]\times[0,\rho_2].$$

Since $-(\partial/(\partial\xi_1))\gamma(\xi)+|(\partial/(\partial\xi_1))\gamma(\xi)|$ and $(\partial/(\partial\xi_2))\gamma(\xi)+|-(\partial/(\partial\xi_2))\gamma(\xi)|$ are zero it is easy to see that

$$|\xi-\eta-h[f(\xi)-f(\eta)]|_1\geqq|\xi-\eta|_1 \quad \text{for all } \xi, \eta\in[0,\rho_1]\times[0,\rho_2]$$

(see Remark 4). Therefore

$$\|\phi-\bar\phi-h[B\phi-B\bar\phi]\|_1\geqq\|\phi-\bar\phi\|_1 \quad \text{for all } \phi, \bar\phi\in D;$$

(iii) is a consequence of (2.15) and Corollary 2 by defining $V[t,\phi]=\|\phi\|_1$ for all $(t,\phi)\in[0,\infty)\times\mathcal{L}^1$. Also, since the off-diagonal entries in the matrix $df(\xi)$ are nonnegative, one may easily check that

$$\lim_{h\to 0^+} d(\xi-\eta+h[f(\xi)-f(\eta)];\mathbb{R}^2_+)/h=0$$

whenever $0\leqq\eta_1\leqq\xi_1\leqq\rho_1$ and $0\leqq\eta_2\leqq\xi_2\leqq\rho_2$. From this it follows that if $C_+=\{\phi\in\mathcal{L}^1: \phi\geqq\theta\}$ then

$$\lim_{h\to 0^+} d(\phi-\bar\phi+h[B\phi-B\bar\phi];C_+)/h=0$$

whenever $\phi, \bar\phi\in D$ and $\phi\geqq\bar\phi$. Since $T(t)$: $C_+\to C_+$ for all $t\geqq 0$ as well, it is also the case that (iv) holds (see [11, Prop. 4.1 and 4.2]). Now suppose that $\sigma>0$. To establish the final assertion it is sufficient from (iii) to show that $\|U(t)\theta\|_1\to\infty$ as $t\to\infty$. Since $U(h)\theta\geqq\theta$ for all $h>0$, we have from (i) and (iv) that

$$U(t+h)\theta=U(t)U(h)\theta\geqq U(t)\theta \quad \text{for all } t, h\geqq 0,$$

and so $\|U(t+h)\theta\|_1\geqq\|U(t)\theta\|_1$ for all $t, h\geqq 0$. Therefore $M=\lim_{t\to\infty}\|U(t)\theta\|_1$ exists (where $0\leqq M\leqq\infty$). Suppose for contradiction, that $M<\infty$. Then $\phi=\lim_{t\to\infty}U(t)\theta$ exists in $\mathcal{L}^1$ (see [7, p. 41]) and this implies that $\phi\in D(A)\cap D$ and $A\phi+B\phi=\theta$. Note that if $\phi=(\phi_1,\phi_2)$, then $a_1\phi_1''+a_2\phi_2''\equiv\theta$, and hence $a_1\phi_1'+a_2\phi_2'$ is constant. However, $a_1\phi_1'+a_2\phi_2'\in\mathcal{L}^1$ so $\phi_1'=a_1^{-1}a_2a_2\phi_2'$ and, in particular, $\phi_1'(0)=0$. But this implies that

$$\phi_1(0)=\sigma,\ \phi_1'(0)=0 \quad \text{and} \quad \phi_1''(\tau)=a_1^{-1}\gamma(\phi_1(\tau),\phi_2(\tau))\geqq 0,$$

which is certainly impossible since $\sigma>0$ and ϕ_1 has a finite integral over $[0,\infty)$. Thus $M=\infty$ and the proof of Proposition 3 is complete.

Remark 5. Since the solution u_χ to (2.17) satisfies $u_\chi(t)\in D$ for all $t\geqq 0$, the mild solution (u_1,u_2) to (2.12) identified with u_χ satisfies $0\leqq u_1(t,\tau)\leqq\rho_1$ and $0\leqq u_2(t,\tau)\leqq\rho_2$ for all $t>0$ and $\tau\in[0,\infty)$. In particular, if $\chi=\theta$ the corresponding solution (μ,ν) to (2.11) satisfies $0\leqq\mu(t,\tau)\leqq c_1$ and $0\leqq\nu(t,\tau)\leqq c_2$ for all $t>0$ and $\tau\in[0,\infty)$. Moreover, since $\lim_{t\to\infty}\|u_\theta(t)\|_1=\infty$, it also follows that

$$\int_0^\infty [\mu(t,\tau)+c_2-\nu(t,\tau)]\,d\tau\to\infty \quad \text{as } t\to\infty.$$

Remark 6. Some of the results of this paper were presented at the London Mathematical Society Symposium on Partial Differential Equations, July 1976, Durham, England.

REFERENCES

[1] P. L. T. BRIAN, J. F. HURLEY AND E. H. HASSELTINE, *Penetration theory for gas absorption accompanied by a second order chemical reaction*, AIChE J., 7 (1961), pp. 226–231.

[2] P. V. DANCKWERTS, *Absorption by simultaneous diffusion and chemical reaction*, Trans. Faraday Soc., 46 (1950), pp. 300–304.

[3] A. FRIEDMAN, *Partial Differential Equations of Parabolic Type*, Prentice-Hall, Englewood Cliffs, NJ, 1964.

[4] ———, *Partial Differential Equations*, Holt, Rinehart and Winston, New York, 1959.

[5] M. IANNELLI, *On a certain class of semilinear evolution systems*, J. Math. Anal. Appl., 56 (1976), pp. 357–367.

[6] C. S. KAHANE, *On a system of nonlinear parabolic equations arising in chemical engineering*, Ibid., 53 (1976), pp. 343–358.

[7] M. A. KRASNOSELSKII, *Positive Solutions of Operator Equations*, Noordhoff, Gronigen, the Netherlands, 1964.

[8] V. LAKSHMIKANTHAM AND S. LEELA, *Differential and Integral Inequalities*, vol. I, Academic Press, New York, 1969.

[9] R. MARTIN, *Invariant sets for perturbed semigroups of linear operators*, Ann. Mat. Pura Appl., 150 (1975), pp. 221–239.

[10] ———, *Nonlinear perturbations of uncoupled systems of elliptic operators*, Math. Ann., 211 (1974), pp. 155–169.

[11] ———, *Nonlinear perturbations of linear evolution systems*, J. Math. Soc. Japan, 29 (1977), pp. 233–252.

[12] A. PAZY, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Lecture Notes, No. 10, Univ. of Maryland, College Park, MD, 1974.

[13] ——— *A class of semilinear equations of evolution*, Israel J. Math., 20 (1975), pp. 23–36.

[14] J. R. A. PEARSON, *Diffusion of one substance into a semi-infinite medium containing another with a second order reaction*, Appl. Sci. Res. Section A, 11 (1963), pp. 321–340.

SIAM J. MATH. ANAL.
Vol. 9, No. 6, December 1978
© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0011 $01.00/0

A RUNGE APPROXIMATION AND UNIQUE CONTINUATION THEOREM FOR PSEUDOPARABOLIC EQUATIONS*

WILLIAM RUNDELL† AND MICHAEL STECHER†

Abstract. In this paper we construct an integral operator for solutions of the pseudoparabolic equation $\Delta_n u_t - A(r)u_t + B(r)u = f$. This is then used to obtain a Runge approximation and unique continuation result for pseudoparabolic equations.

Introduction. In this paper we shall study some properties of the solutions to the equation

$$Lu = \Delta_n u_t - A(r)u_t + B(r)u = f \tag{0.1}$$

where Δ_n denotes the n-dimensional Laplacian operator.

Such equations, usually referred to as of pseudoparabolic or Sobolev type, serve as models in numerous physical applications. This aspect is well documented in the literature, cf. [7], [8].

We shall show that there exists a one to one, onto map between the class of solutions to equation (0.1) and solutions to the equation $\Delta_n h_t = g$. This result was also the subject of [1], [6] under the assumption that the coefficients A and B were analytic functions of $r^2 = x_1^2 + \cdots + x_n^2$. The first section of this paper provides an extension to the case when the coefficients A and B are continuously differentiable functions of the variable r. As in these previous papers the map takes the form of an integral operator, the kernel of which is independent of the domain under consideration, depending only on the coefficients A and B that appear in (0.1). This property is used in later sections of the paper; in the second section to obtain a Runge approximation property and a unique continuation theorem, and in the final section to derive a reflection principle for equations of type (0.1) defined in a cylindrical domain in $R^n \times R$. The unique continuation theorem is interesting since it provides a contrast to the unique continuation properties of parabolic equations.

In conclusion we wish to note that all of the results in this paper also hold for solutions of the following equation:

$$\tilde{L}[u] = \Delta_n u_t - A(r)u_t + \eta\Delta_n u + B(r)u = f, \tag{0.2}$$

where η is a constant. This follows from the fact that the change of dependent variables $v = e^{\eta t}u$ transforms equation (0.1) to (0.2).

1. An integral operator. In this section we construct an integral operator which maps solutions of $\Delta_n h_t = g$ onto solutions of

$$L[u] \equiv \Delta_n u_t - A(r)u_t + B(r)u = f, \tag{1.1}$$

where Δ_n is the n-dimensional Laplacian ($n \geqq 2$), u_t denotes differentiation with respect to t, and $A(r)$, $B(r)$ are continuously differentiable functions of r ($r^2 = \sum_{i=1}^n x_i^2$) for $0 \leqq r \leqq R$.

We will show that solutions of (1.1) may be written in the following form:

$$\begin{aligned} u(\mathbf{x}, t) &= (I + \mathcal{G})[h] \\ &= h(\mathbf{x}, t) + \int_0^t \int_0^1 \sigma^{n-1} G_t(r, 1-\sigma^2, t-\tau) h_t(\mathbf{x}\sigma^2, \tau)\, d\sigma\, d\tau, \end{aligned} \tag{1.2}$$

* Received by the editors July 21, 1976, and in final revised form June 6, 1977.
† Department of Mathematics, Texas A & M University, College Station, Texas 77843.

where $\Delta_n h_t = g$ and g is given in terms of f, and $u(\mathbf{x}, 0)$. In a recent paper on elliptic equations [3], Gilbert constructed a similar operator where the coefficients depended on r^2 but analyticity was not required.

Our method of presentation is to describe certain conditions which the G function, appearing in (1.2), must satisfy. We show such a function exists and then establish our representation (1.2).

In particular, we wish G to satisfy the following,

(1.3a) $$2(1-s)G_{rstt} - G_{rtt} + r(G_{rrtt} - A(r)G_{tt} + B(r)G_t) = 0, \qquad 0<r<R, \quad 0<s<1, \quad \text{and} \quad 0<t,$$

(1.3b) $$G(0, s, t) = 0, \qquad 0 \leqq s \leqq 1, \quad 0 \leqq t,$$

(1.3c) $$G_{rtt}(r, 0, t) = -rB(r), \qquad 0<r<R, \quad 0<t,$$

(1.3d) $$G(r, s, 0) = 0, \qquad 0 \leqq r \leqq R, \quad 0 \leqq s \leqq 1,$$

(1.3e) $$G_t(r, s, 0) = H(r, s), \qquad 0<r<R, \quad 0<s<1.$$

The function $H(r, s)$ appearing in (1.3e) is not arbitrary. It is the solution to the following problem:

(1.4a) $$2(1-s)H_{rs} - H_r + r(H_{rr} - A(r)H) = 0, \qquad 0<r<R, \quad 0<s<1,$$

(1.4b) $$H(0, s) = 0, \qquad 0 \leqq s \leqq 1,$$

(1.4c) $$H_r(r, 0) = rA(r),$$

i.e., $H(r, s)$ is Gilbert's G-function for equation (1.4a) cf. [4]. Thus not only is G a solution to a singular partial differential equation, but some of its initial data is also described as the solution to a singular Goursat problem.

It is easily seen that solutions of (1.3) and (1.4) exist for $0 \leqq s < 1$. The difficulty lies in obtaining bounds on G and some of its partial derivatives as s approaches one. We first show that a continuous solution of (1.4) exists for s in the interval [0, 1].

Making the change of variables

(1.5) $$\rho = r\sqrt{1-s}, \qquad \eta = s,$$

in equation (1.4a), rewriting (1.4) as an integral equation, and then returning to the original coordinates, we see that $H(r, s)$ is a solution of (1.4) if and only if H satisfies the following equation:

(1.6) $$H(r, s) = (1-s)^{-1} \int_0^s \int_0^{r\sqrt{(1-s)/(1-z)}} \frac{y}{2} A(y)H(y, z)\, dy\, dz + (1-s)^{-1} \int_0^{r\sqrt{1-s}} yA(y)\, dy, \qquad 0 \leqq r \leqq R, \quad 0 \leqq s < 1.$$

We define the operator T as follows,

(1.7) $$T[f](r, s) \equiv (1-s)^{-1} \int_0^s \int_0^{r\sqrt{(1-s)/(1-z)}} \frac{y}{2} A(y)f(y, z)\, dz\, dz,$$

where f is a continuous function defined on $[0, R] \times [0, 1]$. From (1.6) it is clear that

(1.8) $$H(r, s) = \sum_{n=0}^{\infty} T^n[F](r, s),$$

where

$$F(r,s)=(1-s)^{-1}\int_0^{r\sqrt{1-s}} yA(y)\,dy. \tag{1.9}$$

Routine estimates give us the following theorem:

THEOREM 1.1 *The solution $H(r,s)$ of* (1.4) *is continuous on* $[0,R]\times[0,1]$, *twice continuously differentiable for* $0\leqq s<1$, *and satisfies the following inequalities for some constant c,*

$$\begin{aligned} &|H(r,s)|\leqq cr^2, \qquad |H_s(r,s)|\leqq\frac{c}{1-s}, \qquad |H_r(r,s)|\leqq cr,\\ &|H_{rs}(r,s)|\leqq\frac{c}{1-s}, \qquad |H_{rr}(r,s)|\leqq c. \end{aligned} \tag{1.10}$$

We now return to problem (1.3) and show that the solution G is bounded as in (1.10). Using the variables (1.5) we transform (1.3) into the following equivalent integral equation.

$$\begin{aligned} G(r,s,t)=F(r,s,t)&+(1-s)^{-1}\int_0^s\int_0^{r\sqrt{(1-s)/(1-z)}}\frac{y}{2}A(y)G(y,z,t)\,dy\,dz\\ &-(1-s)^{-1}\int_0^s\int_0^{r\sqrt{(1-s)/(1-z)}}\int_0^t\frac{y}{2}B(y)G(y,z,\tau)\,d\tau\,dy\,dz, \end{aligned} \tag{1.11}$$

where $F(r,s,t)$ is defined by

$$\begin{aligned} F(r,s,t)=tH(r,s)&-\frac{t^2}{2(1-s)}\int_0^{r\sqrt{1-s}} yB(y)\,dy\\ &-\frac{t}{(1-s)}\int_0^s\int_0^{r\sqrt{(1-s)/(1-z)}}\frac{y}{2}A(y)H(y,z)\,dy\,dz. \end{aligned}$$

Observing that the integral equation (1.11) is of the same form as (1.6), we have the following theorem:

THEOREM 1.2. *Let $A(r)$ and $B(r)$ be continuously differentiable on* $[0,R]$. *Then there exists a solution, $G(r,s,t)$, of* (2.3) *which is continuous on* $[0,R]\times[0,1]$, *twice continuously differentiable for* $0\leqq s<1$, *and for some constant C we have the following inequalities,*

$$\begin{aligned} &|G_r(r,s,t)|\leqq Cr, \qquad |G_s(r,s,t)|\leqq\frac{C}{1-s},\\ &|G_{rs}(r,s,t)|\leqq\frac{C}{1-s}, \qquad |H_{rr}(r,s,t)|\leqq C, \end{aligned} \tag{1.12}$$

for $0\leqq r\leqq R$ *and* $0\leqq s<1$.

We remark that G_t also satisfies inequalities (1.12). To see this one differentiates equation (1.11) with respect to t, and observes G_t satisfies an equation of the same form.

We now proceed to show that the operator $1+\mathcal{G}$ defined by (1.2) maps solutions of $\Delta_n h_t=g\,(n\geqq 2)$ into solutions of (1.1). If we define $h(\mathbf{x},0)$ to be equal to $u(\mathbf{x},0)$ and then determine $h_t(\mathbf{x}, t)$ uniquely from $u_t(\mathbf{x}, t)$ we will have shown that (1.2) is invertible. Thus we differentiate (1.2) with respect to t and obtain

$$(1.13)\qquad \begin{aligned} u_t(\mathbf{x}, t) = h_t(\mathbf{x}, t) + \int_0^1 \sigma^{n-1} G_t(r, 1-\sigma^2, 0) h_t(\mathbf{x}\sigma^2, t)\, d\sigma \\ + \int_0^t \int_0^1 \sigma^{n-1} G_{tt}(r, 1-\sigma^2, t-\tau) h_t(\mathbf{x}\sigma^2, \tau)\, d\sigma\, d\tau. \end{aligned}$$

Using the change of variables

$$\mathbf{x} = (r, \boldsymbol{\theta}), \qquad \sigma^2 = \frac{\rho}{r},$$

where $(r, \boldsymbol{\theta})$ are the n-dimensional spherial coordinates of $\mathbf{x}$, it has been shown in [1] that (1.13) is invertible. Hence equation (1.2) is invertible.

The invertibility of (1.2) by itself is of course not significant. What is important is the fact that $h(\mathbf{x}, t)$ is a solution of the simpler partial differential equation

$$(1.14)\qquad \Delta_n h_t = g(x, t), \qquad h(\mathbf{x}, 0) = u(\mathbf{x}, 0).$$

In order to see this we will compute $L[u]$.

$$(1.15)\qquad \begin{aligned} f(\mathbf{x}, t) &= L[u] \\ &= \Delta_n h_t + B(r) h(\mathbf{x}, 0) + \int_0^1 \sigma^{n+3} H(r, 1-\sigma^2) \Delta_n h_t(\mathbf{x}\sigma^2, t)\, d\sigma \\ &\quad + \int_0^t \int_0^1 \sigma^{n+3} G_{tt}(r, 1-\sigma^2, t-\tau) \Delta_n h_t(\mathbf{x}\sigma^2, \tau)\, d\sigma\, d\tau. \end{aligned}$$

Arguing as in (1.13), we see that $\Delta_n h_t = g$, where $g(x, t)$ is the solution to the integral equation

$$(1.16)\qquad \begin{aligned} f(\mathbf{x}, t) - B(r) u(\mathbf{x}, 0) = g + \int_0^1 \sigma^{n+3} H(r, 1-\sigma^2) g(\mathbf{x}\sigma^2, t)\, d\sigma \\ + \int_0^t \int_0^1 \sigma^{n+3} G_{tt}(r, 1-\sigma^2, t-\tau) g(\mathbf{x}\sigma^2, \tau)\, d\sigma\, d\tau. \end{aligned}$$

In conclusion, what we have established is the following. Every smooth solution $u(\mathbf{x}, \mathrm{t})$ of (1.1) has the following representation,

$$(1.2)\qquad u(\mathbf{x}, t) = h(\mathbf{x}, t) + \int_0^t \int_0^1 \sigma^{n-1} G_t(r, 1-\sigma^2, t-\tau) h_t(\mathbf{x}\sigma^2, \tau)\, d\sigma\, d\tau,$$

where $h(\mathbf{x}, t)$ satisfies

$$\Delta_n h_t = g(x, t),$$

and g satisfies (1.16).

2. Runge approximation and unique continuation properties. It is well known that solutions of Laplace's equation satisfy the Runge approximation property. More precisely, if u is a solution of $\Delta u = 0$ in domain D and D^* is a domain such that $D \subset D^*$ and

(R) the set D^*/D has no component relatively compact in D^*, then for every compact set $K \subset D$ and $\varepsilon > 0$ there is a solution v of Laplace's equation in D^* such that $\sup_{x \in K} |v(x) - u(x)| < \varepsilon$.

The following theorem shows that a similar result holds for solutions of pseudoparabolic equations.

THEOREM 2.1 (Runge approximation theorem). *Let D and D^* ($D \subset D^*$) be two open connected subsets of $R^n (n \geqq 2)$ which satisfy condition* (R), *and let T be an interval of the real line. Let A and B be continuously differentiable functions of r in D^* and $f(x, t)$ be continuous in $D^* \times T$. Then every solution $u(x, t)$ of*

$$Lu = \Delta_n u_t - Au_t + Bu = f \tag{2.1}$$

in $D \times T$ may be approximated uniformly on compact subsets by a solution $u^*(x, t)$ defined in the larger domain $D^* \times T$.

Proof. The proof relies on the fact that the kernel of the integral operator $G(r, \sigma, t-\tau)$ depends only on the coefficients and is independent of the domain. Thus it is sufficient to show that the theorem holds for solutions of the simple pseudoparabolic equation

$$\Delta_n h_t = g. \tag{2.2}$$

We note from (1.16) that g is defined and continuous on $D^* \times T$. By linearity we may assume that $h(x, 0) = 0$.

To see that (2.2) has the Runge property, one uses the Runge property for Poisson's equation to approximate h_t at a finite number of times t_i. Now connect these approximates together in a linear manner and integrate from 0 to t.

A similar approach allows us to prove a unique continuation theorem for solutions of the homogeneous version of (2.1), namely,

THEOREM 2.2 (Unique continuation theorem). *Let D and D^* be connected open subsets of $R^n (n \geqq 2)$ with $D \subset D^*$ and let T be an interval of the real line. Let*

$$S(D^* \times T) = \{u(x, t): \Delta u_t \in C^0(D^* \times T), Lu = 0 \text{ and } u(x, 0) = 0, \text{for } x \in D^*\}.$$

Then if $u \in S(D^* \times T)$ such that $u = 0$ in $D \times T$ then $u = 0$ in $D^* \times T$.

Proof. Denote by u the element of $S(D^* \times T)$ such that $u = 0$ in $D \times T$. Then u has the representation

$$u = (I + \mathcal{G})h, \tag{2.3}$$

where h is a solution of $h_t = 0$ with $h(x, 0) = 0$. By the uniqueness of the representation (2.3), $h(x, t)$, and hence $h_t(x, t)$ must vanish identically in $D \times T$. By the unique continuation theorem for Laplace's equation it follows that $h_t(x, t)$ must vanish in all of $D^* \times T$, and since $h(x, 0) = 0$ we have $h(x, t) \equiv 0$ in $D^* \times T$. The map (2.3) now gives $u(x, t) \equiv 0$ in $D^* \times T$, and the proof is complete.

We note that the above result is true for solutions of the heat equation without the added assumption that $u(x, 0) = 0$. In the case of one space variable this follows from the observation that for some x_0 and $\delta > 0$ we have $u(x, t) = 0$ in the strip $D_0 \times T = \{x_0 - \delta < x < x_0 + \delta, t \in T\}$, and hence $u(x_0, t) = u_x(x_0, t) = 0$, $t \in T$. Thus by the uniqueness to the Cauchy problem for a noncharacteristic line we have $u(x, t) \equiv 0$ in any connected region of $R \times T$ containing $D_0 \times T$ in which u is a solution of the heat equation.

For the pseudoparabolic equation, however, the assumption that $u(x, 0) = 0$ (or at least that $u(x, 0)$ be specified a priori) is essential to Theorem 2.2. The simplest illustration of this is for solutions of the equation $h_{xxt} = 0$, for example

$$h(x, t) = \begin{cases} 0 & x < 0, \\ e^{-1/x} & x \geqq 0. \end{cases}$$

Theorems 2.1 and 2.2 are also true in the one dimensional case. This can be seen by using the methods of [2].

3. A reflection principle. We give a further illustration of the use of the integral operator developed in § 1 by using it to obtain a reflection principle for solutions of the equation $Lu=0$.

Our purpose will not be to state the most general result possible but to illustrate the technique. One can clearly generalize our result; for example to the case of time dependent coefficients by the use of the integral operator developed in [6], to non-homogeneous boundary conditions and analytic surfaces rather than hyperplanes.

THEOREM 3.1. *Let $D\times T$ be a connected cylindrical domain in the half space $x_1>0$ whose boundary contains a portion $\sigma\times T$ of the hyperplane $x_1=0$. Denote by $\tilde{D}$ the domain formed by reflecting D about the hyperplane $x_1=0$. Suppose that $A(r^2)$, $B(r^2)$ and $u(x,0)$ are analytic functions of their independent variables in the domain $D\cup\sigma\cup\tilde{D}$. Let $u(x,t)$ be a solution of $Lu=0$ in $D\times T$ such that $u=0$ on $\sigma\times T$; then $u(x,t)$ can be uniquely continued as a solution into all of $(D\cup\sigma\cup\tilde{D})\times T$.*

Proof. By linearity we can assume that $u(x,0)=0$ and that u satisfies $Lu=f$ in $D\times T$ where $f=L\{u(x,0)\}$ is an analytic function defined in all of $D\cup\sigma\cup\tilde{D}$. Since the coefficients A and B are analytic functions, so also is the kernel $G(r^2,1-\sigma^2,t-\tau)$ with respect to r^2 and σ. Thus there exists a function $h(x,t)$ such that

$$u=h+\int_0^t\int_0^1 \sigma^{n-1}G_t(r^2,1-\sigma^2,t-\tau)h_\tau(x\sigma^2,\tau)\,d\sigma\,d\tau, \tag{3.1}$$

and h satisfies the equation $\Delta h_t=g$ where g is given in terms of f and G by equation (1.16). Thus since f, A and B are defined and analytic in all of $D\cup\sigma\cup\tilde{D}$ so also is $g(x,t)$ as a function of x. We also have $h(x,0)=0$. Since $u(x,t)=0$ on σ we have $h(x,t)=0$ on σ (the Volterra equation (3.1) has a unique solution).

Thus by the reflection principle for solutions of Poisson's equation $h(x,t)$ may be extended as a solution of $\Delta h_t(x,t)=g$ with $h(x,0)=0$ to all of $(D\cup\sigma\cup\hat{D})\times T$. By substituting this extension into (3.1), the proof is complete.

REFERENCES

[1] D. L. COLTON, *Integral operators and the first initial boundary value problems for pseudoparabolic equations with analytic coefficients*, J. Differential Equations, 13 (1973), pp. 507–520.

[2] ———, *Pseudoparabolic equations in one space variable*, Ibid., 12 (1972), pp. 559–565.

[3] R. P. GILBERT, *Integral operator methods for approximating solutions of Dirichlet problems*, ISNM, 15 (1970), pp. 129–146.

[4] ———, *The construction of boundary value problems by function theoretic methods*, this Journal, 1 (1970), pp. 96–114.

[5] P. D. LAX, *A stability theory for solutions of abstract differential equations, and its application to the study of the behavior of solutions of elliptic equations*, Comm. Pure Appl. Math., 9 (1956), pp. 747–766.

[6] W. RUNDELL AND M. STECHER, *A method of ascent for parabolic and pseudoparabolic partial differential equations*, this Journal, 7 (1976), pp. 898–912.

[7] R. E. SHOWALTER, *The Sobolev equation* I. Applicable Anal., 5 (1975), pp. 15–22.

[8] ———, *The Sobolev equation II*, Applicable Anal., 5 (1975), pp. 81–99.

SIAM J. MATH. ANAL.
Vol. 9, No. 6, December 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0012 $01.00/0

ESTIMATES FOR THE GREEN'S FUNCTIONS OF ELLIPTIC OPERATORS*

CATHERINE BANDLE†

Abstract. Inequalities for the Green's functions of $Lu = -\Delta u - pu$ are derived by means of the level line technique and bounds for linear boundary value problems are constructed.

Introduction. In this paper we study the Green's function of the operator $\Delta + p(x)$ in a bounded domain. First we derive some comparison theorems based on a maximum principle, and then we prove a differential inequality for a certain expression involving the Green's function. By exhibiting this inequality, we obtain estimates for the Green's functions which include as a special case Weinberger's bounds [16] for their norms. Moreover, it is possible to construct upper and lower bounds for the solutions of boundary value problems and to generalize results of [1]. This work was motivated by a theorem of Pólya and Szegö [14] concerning the warping function. If D is a plane domain and $\dot{R}$ its maximal conformal radius, then the theorem says that the solution of $\Delta u + 1 = 0$ in D, $u = 0$ on ∂D satisfies $\max u(x) \geqq \dot{R}^2/4$. Payne [13] extended this result to higher dimensions by using the Green's function of the Laplace operator. His arguments together with a geometrical inequality led to similar statements for the problem $\Delta u + e^u = 0$ in D, $u = 0$ on ∂D [3]. The question arose whether the solutions of a more general problem of the type $\Delta u + f(u) = 0$ in D, $u = 0$ on ∂D could be estimated by means of the conformal radius. A partial answer is given in this paper for $f(u) = \alpha u + 1$ in the last section.

1. The Green's function.

1.1. Definition. Let $D \subset R^N$, $N = 2, 3$ be a bounded domain with a smooth boundary, and let $x = (x_1, \cdots, x_N)$ denote an arbitrary point in R^N. Consider the differential operator

$$Lu = -\Delta u - p(x)u,$$

where Δ stands for the Laplacian and $p(x)$ is a real analytic function in D of class $C^0(\bar{D})$.[1] Let us introduce the function

$$s(x, y) = \begin{cases} \dfrac{1}{2\pi} \log |x - y|^{-1} & \text{if } N = 2, \\ \dfrac{1}{4\pi} |x - y|^{-1} & \text{if } N = 3 \end{cases}$$

which is a fundamental solution of the Laplace equation. The Green's function $g(x, y)$ of L with respect to D is defined by the following requirements [5], [8].

(i) For fixed $y \in D$, $h(x, y) = g(x, y) - s(x, y)$ is a continuous function in $\bar{D}$.

(ii) For fixed $y \in D$, h solves $Lh = p(x)s(x, y)$ in D.

(iii) $g(x, y) = 0$ for $x \in \partial D$, $y \in D$.

According to the classical theory, $g(x, y)$ exists and is uniquely determined whenever the boundary value problem $Lv = 0$ in D, $v = 0$ on ∂D has only the trivial solution. We observe that $Lg(x, y) = \delta_y$ in D, δ_y being the Dirac measure at the point y. The importance of the Green's function is that every classical solution of the problem

* Received by the editors January 28, 1977, and in revised form May 13, 1977.

† Mathematisches Institut der Universität Basel, Basel, Switzerland.

[1] This assumption has been made to avoid lengthy technical discussions which could arise in § 2 in connection with questions of regularity.

$Lu = f(x)$ in D with $u = \nu(x)$ on ∂D has the representation

$$u(x) = \int_D g(x, y) f(y)\, dy - \oint_{\partial D} \frac{\partial g(x, y)}{\partial n} \nu(y)\, ds.$$

Here, n is the outer normal, ds is the surface element on ∂D and $dx = dx_1\, dx_2 \cdots dx_N$. In view of the particular form of the singularity of s, the operator

$$Gu = \int_D g(x, y) u(y)\, dy$$

is a mapping from $L^2(D)$ into $C^0(D)$. A detailed study of the Green's function is found in [5] (see also [7]). More specific questions such as its behavior under variation of the domain are treated by Garabedian and Schiffer in [8].

1.2. If $p(x) \leqq 0$, the maximum principle for subharmonic functions implies that $g(x, y)$ is positive in D. Let us now give a criterion for $g(x, y)$ to be positive in the case where p changes sign. We have

$$p(x) = p^+(x) - p^-(x), \quad \text{where } p^\pm(x) = \max\{\pm p(x), 0\}.$$

Consider the eigenvalue problem

$$\Delta\varphi + [\mu p^+(x) - p^-(x)]\varphi = 0 \quad \text{in } D, \qquad \varphi = 0 \quad \text{on } \partial D. \tag{1.1}$$

It has a countable number of eigenvalues $0 < \mu_1 < \mu_2 \leqq \cdots$ provided $p^+(x) \not\equiv 0$. According to the Rayleigh principle

$$\mu_1 = \min_{v=0 \text{ on } \partial D} \frac{\int_D \operatorname{grad}^2 v\, dx + \int_D v^2 p^-\, dx}{\int_D v^2 p^+\, dx}$$

where v ranges among all piecewise differentiable functions. μ_1 can also be characterized by [10]

$$\mu_1 = \sup_v \min_{x \in \bar{D}} \frac{-\Delta v + p^-(x) v}{p^+(x) v}, \qquad v \not\equiv 0 \text{ in } \operatorname{supp} p^+(x) \tag{1.2}$$

where v ranges over all positive functions of class $C^2(D)$. The following result was known to different authors [4], and others, and it will be repeated for the sake of completeness.

LEMMA 1.1. *The Green's function $g(x, y)$ is positive in D if and only if $\mu_1 > 1$.*

Proof. (i) Let $\mu_1 > 1$. The corresponding eigenfunction $\varphi_1(x)$ is of constant sign and can be taken positive in D. Assume that $g(x, y)$ is negative in a domain $D^- \subseteq D$. Since $g(x, y)$ behaves like $s(x, y)$ near the point y, D^- does not contain y. The function $g(x, y)$ is therefore a solution of (1.1) with D replaced by D^-. The corresponding eigenvalue is $\mu = 1$. Since $g(x, y)$ does not change sign in D^-, $\mu = 1$ is the smallest eigenvalue. Because of the monotonicity of the eigenvalues with respect to the domain D [6], it follows that $\mu_1(D) \leqq \mu_1(D^-) = 1$, which contradicts our assumption. By Hopf's maximum principle g cannot take its minimum at an interior point. Hence, $g(x, y) > 0$ in D.

(ii) Let $g(x, y)$ be positive in D. Then the solution of $Lw = 1$ in D, $w = 0$ on ∂D is positive in D. By (1.2) we have

$$\mu_1 > \min \frac{p^+(x) w + 1}{p^+(x) w} > 1,$$

which completes the proof of Lemma 1.1.

As an immediate consequence we find

COROLLARY 1.1. *If the Green's function is positive in D, then every solution of* $Lw \geqq 0$ *in D,* $w \geqq 0$ *on* ∂D *is positive in D unless* $w \equiv 0$.

Next, we describe some simple properties of $g(x, y)$, all based on Corollary 1.1. They illustrate the behavior of $g(x, y : p)$ as a function of p.

(A) *If* $g(x, y : p) \geqq 0$, *then* $g(x, y : q) \geqq 0$ *for all* $q \leqq p$.

Proof. Since $p^+ \geqq q^+$ and $p^- \leqq q^-$, we conclude from the Rayleigh principle that $\mu_1(q^+, q^-) \geqq \mu_1(p^+, p^-) > 1$. (A) follows now from Lemma 1.1.

(B) *If* $p_1(x) \geqq p_2(x)$ *and* $g(x, y : p_1) \geqq 0$, *then* $g(x, y : p_1) \geqq g(x, y : p_2)$.

Proof. The difference $d(x) = g(x, y : p_1) - g(x, y : p_2)$ satisfies $\Delta d + p_1 d \leqq 0$ in D, $d = 0$ on ∂D. In view of Corollary 1.1, $d(x) \geqq 0$.

(C) *If* $p_i(x) \geqq 0$ *for* $i = 1, 2$, $p = p_1 + p_2$ *and* $g(x, y : p) \geqq 0$, *then* $g(x, y : p_1) + g(x, y : p_2) - g(x, y : p) \leqq g(x, y : 0)$.

Proof. Consider the function $d(x) = g(x, y : p_1) + g(x, y : p_2) - g(x, y : p)$. We have

$$\Delta d + p_1[g(x, y : p_1) - g(x, y : p)] + p_2[g(x, y : p_2) - g(x, y : p)] = -\delta_y$$

and because of our assumptions and (B)

$$\Delta d \geqq -\delta_y.$$

Hence,

$$\Delta d - \Delta g(x, y : 0) \geqq 0.$$

The maximum principle for harmonic functions yields the desired result.

Similarly we show the next property.

(D) *If* $p_i(x) \geqq 0$ *for* $i = 1, 2$, $p = p_1 - p_2$ *and* $g(x, y : p_1) \geqq 0$, *then* $g(x, y : p_1) + g(x, y : -p_2) - g(x, y : p) \geqq g(x, y : 0)$.

In particular

$$g(x, y : p^+) + g(x, y : -p^-) - g(x, y : p) \geqq g(x, y : 0).$$

This type of inequalities has already been established by Luttinger [12] for the Green's function associated with the Schrödinger equation.

Consider a function $p(x, t)$ which depends on x and a real variable t.

LEMMA 1.2. *If* $p(x, t)$ *is twice continuously differentiable in t, then* $g(x, y : t) \equiv g(x, y : p(x, t))$ *is a differentiable function of t. Moreover,* $\partial g/\partial t \equiv \dot{g}$ *satisfies* $\Delta \dot{g} + p(x, t)\dot{g} + \dot{p}(x, t) g = 0$ *in D and* $\dot{g} = 0$ *on* ∂D.

Proof. The difference $dg = g(x, y : t + \Delta t) - g(x, y : t)$ can be written in the form

$$dg = \int_D g(x, z : t) g(y, z : t + \Delta t)[\dot{p}(z, t)\Delta t + \tfrac{1}{2}\ddot{p}(z, \tau)(\Delta t)^2]\, dz \qquad (t \leqq \tau \leqq t + \Delta t).$$

Hence

$$\dot{g}(x, y : t) = \lim_{\Delta t \to 0} \int_D g(x, z : t) g(y, z : t + \Delta t) \dot{p}(z, t)\, dz.$$

Recalling that $g(y, z : t + \Delta t) = s(y, z) + h(y, z : t + \Delta t)$ where h is bounded, we can apply Lebesgue's dominated convergence theorem, and we find

$$\dot{g}(x, y : t) = \int_D g(x, z : t) g(y, z : t) \dot{p}(z, t)\, dz$$

which is equivalent to our assertion.

(E) *Let $p(x,t)$ be a three times continuously differentiable function of t such that $\dot p(x,t)$ is of one sign and $\ddot p(x,t)\geqq 0$. If $g(x,y:t)\geqq 0$ then $g(x,y:t)$ is convex in t.*

Proof. As in the proof of Lemma 1.2 we show that $\ddot g(x,y:t)$ exists and solves $\Delta\ddot g+p\ddot g+2\dot p\dot g+\ddot p g=0$ in D, $\ddot g=0$ on ∂D. According to Lemma 1.2 and Corollary 1.1, $\dot p\dot g\geqq 0$. We have therefore $L\ddot g\geqq 0$ and by Corollary 1.1, $\ddot g\geqq 0$.

We shall add a last property observed by Garabedian and Schiffer [8] which is readily obtained by means of Corollary 1.1.

(F) *If $g(x,y)$ and $\tilde g(x,y)$ are the Green's functions in D or $\tilde D$, respectively, and if $\tilde D\subseteq D$, then $g(x,y)\geqq\tilde g(x,y)$ in $\tilde D$ provided that $g(x,y)\geqq 0$.*

Remark. All statements (A)–(D) hold also for the solutions of $Lw=f$ in D, $w=\nu$ on ∂D, provided f and ν are nonnegative. This follows immediately from the integral representation of w involving the Green's function.

2. A differential inequality for Green's functions

2.1. In this section we shall use the following notations.

$$D(\mu)=\{x\in D\colon g(x,y)>\mu\},\qquad \Gamma(\mu)=\partial D(\mu),\qquad a(\mu)=\int_{D(\mu)}dx,\qquad \mu_0\leqq\mu\leqq\infty.$$

Because of the analyticity of $g(x,y)$ in $D-\{y\}$, $a(\mu)$ is a continuous, decreasing function with $a(\mu_0)=A$ and $a(\infty)=0$. With $\mu(a)$ we shall denote the inverse of $a(\mu)$. It corresponds to the value of g on the level surface enclosing a body of a volume a, and satisfies locally a Lipschitz condition [2]. Furthermore we introduce the concept of symmetrization.

DEFINITION. If $B\subset R^N$ is a bounded domain, then the symmetrized domain B^* is the sphere $\{x\colon |x|<\rho\}$ with the same volume as B. (If B is compact, then we take for B^* the closed sphere.)

DEFINITION. Let $u(x)$ be a real function in D, and $B(\mu)=\{x\in D\colon u(x)\geqq\mu\}$. The symmetrized function $u^*\colon D^*\to R$ is defined as follows:

$$u^*(x)=\sup\{\mu\colon x\in B^*(\mu)\}.$$

Since

$$\left(\frac{d\mu}{da}\right)^{-1}=-\oint_{\Gamma(\mu)}\frac{ds}{|\operatorname{grad} g|},$$

it follows from Schwarz's inequality that

$$-\frac{1}{(d\mu/da)}\oint_{\Gamma(\mu)}|\operatorname{grad} g|\,ds\geqq\left\{\oint_{\Gamma(\mu)}ds\right\}^2. \tag{2.1}$$

We observe that $\oint_{\Gamma(\mu)}|\operatorname{grad} g|\,ds=\int_{D(\mu)}pg\,dx+1$, and that according to the geometrical isoperimetric inequality

$$\left\{\oint_{\Gamma(\mu)}ds\right\}^2\geqq q(a)=\begin{cases}4\pi a & \text{if } N=2,\\ (36\pi a^2)^{2/3} & \text{if } N=3.\end{cases}$$

Hence

$$\int_{D(\mu)}pg\,dx+1\geqq -q(a)\frac{d\mu}{da}. \tag{2.2}$$

In order to evaluate the left-hand side of this inequality we recall the rearrangement theorem in [9], namely

$$\int_{D(\mu)} p(x)g(x, y)\, dx \leqq \int_{D^*(\mu)} p^*(x)g^*(x, y)\, dx = \int_0^a \rho(a')\mu(a')\, da',$$

where

$$\rho(\omega_N|x|^N) = p^*(x) \quad \text{and} \quad \omega_N = \begin{cases} \pi & \text{if } N=2, \\ (\frac{4}{3})\pi & \text{if } N=3. \end{cases}$$

DEFINITION. $H(a) = \int_0^a \rho(a')\mu(a')\, da' + 1$.

Inserting the previous estimate into (2.2) we find

$$q(a)\frac{d\mu}{da} + H(a) \geqq 0. \tag{2.3}$$

2.2. Let us first discuss the inequality (2.3) in the case where $\rho(a) \equiv 0$, that is $p \equiv 0$. Then it becomes

$$\frac{d\mu}{da} \geqq -1/q(a).$$

Integration from a to A yields

$$\mu(a) \leqq \begin{cases} (4\pi)^{-1}\{\log a^{-1} - \log A^{-1}\} & \text{if } N=2, \\ (3\omega_3^{2/3})^{-1}\{a^{-1/3} - A^{-1/3}\} & \text{if } N=3. \end{cases} \tag{2.4}$$

Integration from 0 to a yields

$$\mu(a) \geqq \begin{cases} (4\pi)^{-1}\{\log a^{-1} - \log (\pi R_y^2)^{-1}\} & \text{if } N=2, \\ (3\omega_3^{2/3})^{-1}\{a^{-1/3} - \omega_3^{-1/3} R_y^{-1}\} & \text{if } N=3, \end{cases} \tag{2.5}$$

where the quantity R_y is defined by the relation

$$\lim_{x \to y} [g(x, y) - s(x, y)] = \begin{cases} (1/(2\pi)) \log R_y & \text{if } N=2, \\ -(1/(3\omega_3))R_y^{-1} & \text{if } N=3. \end{cases}$$

Remark. If D is simply connected and $N=2$, then R_y corresponds to the conformal radius [14]. For $N=3$, R_y is called the harmonic radius [11].

2.3. Let now $\rho(a) \neq 0$ in $(0, A)$. The derivative $d\mu/da$ can then be expressed by $(H'/\rho)'$, and (2.3) becomes

$$q(a)\left(\frac{H'}{\rho}\right)' + H \geqq 0 \quad \text{in } (0, A). \tag{2.6}$$

In order to discuss this inequality we shall introduce the Green's function $G(x, y)$ defined in the following way. For fixed $y \in D^*$

$$\Delta G(x, y) + p^*(x)G(x, y) = -\delta_y \quad \text{in } D^*, \qquad G(x, y) = 0 \quad \text{on } \partial D^*.$$

We have

PROPOSITION 2.1. *If $G(x, 0) > 0$ in D^*, then $G(x, y) > 0$ in D^* for all $y \in D^*$.*

Proof. Since this statement is trivial for negative p^*, we can concentrate on the case where $p^* > 0$.

According to Lemma 1.1 it suffices to prove that the smallest eigenvalue of $\Delta\varphi + \mu^* p^* \varphi = 0$ in D^*, $\varphi = 0$ on ∂D^*, say μ_1^*, is greater than 1. Suppose that this is not true. Let $\varphi_1>0$ be the first eigenfunction. Then by the Green's identity we have

$$\int_{D^*}[\varphi_1\,\Delta G(x,0)-\Delta\varphi_1 G(x,0)]\,dx=(\mu_1^*-1)\int_{D^*}\varphi_1 Gp^*\,dx-\varphi_1(0)=0.$$

Since the left-hand side of this identity is negative, we are led to a contradiction.

LEMMA 2.1. *If $p\geqq 0$ in D, then $G(x,0:p^*)>0$ in D^* implies $g(x,y:p)>0$ in D.*

Proof. By Proposition 2.1 we have $\mu_1^*>1$. Applying Schwarz's estimate [15] $\mu_1>\mu_1^*$, we deduce from Lemma 1.1 that $g(x,y:p)$ is positive in D.

PROPOSITION 2.2. *If $G(x,0)$ is positive in D^*, then $G(x,0)$ is a radially symmetric and a decreasing function of $|x|$.* (Its level surfaces are therefore spheres.)

Proof. The symmetry follows from the uniqueness of $G(x,0)$. For negative functions $p^*(x)$, the monotony is a consequence of the maximum principle. Let p^* be positive and assume that $G(x,0)$ is not everywhere decreasing. Then there exist values $r_1<r_2$ such that $G(r_1,0)=G(r_2,0)=\lambda$ and $G(x,0)<\lambda$ in $D'=\{x: r_1<|x|<r_2\}$. In D' the function $w=G(x,0)-\lambda$ satisfies $\Delta w+p^*w\leqq 0$. If we extend this function as zero outside of D', we obtain an admissible function for the variational characterization of μ_1^*. Hence

$$\mu_1^*\leqq\int_{D'}\operatorname{grad}^2 w\,dx\Big/\int_{D'}p^*w^2\,dx<1.$$

On the other hand it follows from Lemma 1.1 and Proposition 2.1 that $\mu_1^*>1$, which leads to a contradiction.

From now on all quantities related with $G(x,0)$ will be denoted by *, such as $H^*(a)$ and $\mu^*(a)$. By the same methods as in § 2.1 we prove that for positive $G(x,0)$

$$q(a)\left(\frac{H^{*\prime}}{\rho}\right)'+H^*=0\quad\text{in }(0,A).\tag{2.7}$$

Here, we have essentially used the fact that the level surfaces of $G(x,0)$ are spheres. If $G(x,0)$ is positive, then

$$H^*(A)=\int_{D^*}p^*(x)G(x,0)\,dx+1=-\int_{D^*}\Delta G(x,0)\,dx=-\oint_{\partial D^*}\frac{\partial G}{\partial n}\,ds>0.$$

$H^*(a)$ is therefore positive regardless of the sign of p^*.

LEMMA 2.2. *If $G(x,0)$ is positive, then $F(a)=H^*H'/\rho-HH^{*\prime}/\rho$ is nondecreasing in $(0,A)$.*

Proof. We multiply (2.6) with H^* ($\geqq 0$) and (2.7) with H and subtract the two expressions. Then

$$H^*\left[\frac{H'}{\rho}\right]'-H\left[\frac{H^{*\prime}}{\rho}\right]'=\left[\frac{H^*H'}{\rho}-\frac{HH^{*\prime}}{\rho}\right]'\geqq 0.$$

From this result we obtain the

THEOREM 2.1. *If $G(x,0)\geqq 0$ and $p(x)\geqq 0$, then for fixed $y\in D$*

$$0\leqq g^*(x,y)\leqq G(x,0).$$

Proof. Assume first that $p(x)>0$. By Lemma 2.2 we have $F(a)\leqq F(A)$, and in view of the positivity of g and G, $H'(A)=H^{*\prime}(A)=0$. This implies $F(a)\leqq 0$ and $F(a)\rho(a)/H^*(a)^2=[H(a)/H^*(a)]'\leqq 0$. Hence, $H(a)\leqq H^*(a)$ in $(0,A)$ and $H'/\rho=\mu(a)\leqq H^{*\prime}/\rho\leqq\mu^*(a)$. Since $\mu(a)|_{a=\omega_N|x|^N}=g^*(x,y)$ and $\mu^*(a)|_{a=\omega_N|x|^N}=G(x,0)$, the assertion is established. If $p(x)\geqq 0$, we apply first the inequality (2.3) in the interval where $\rho(a)$ vanishes. The other part of the proof remains the same.

Remark. If $p(x)\leqq 0$, we can only conclude that

$$H(a)\geqq H^*(a). \tag{2.8}$$

The stronger inequality $\mu(a)\leqq \mu^*(a)$ does not follow from our considerations. It is not clear whether it is true or not.

2.4. This part deals with a generalization of inequality (2.5). Besides the Green's function $g(x, y)$ defined in § 1.1 we shall consider the solution of

$$\Delta\tilde{g}(x, 0:R)+p^*(x)\tilde{g}(x, 0:R)=-\delta_0 \quad \text{in } \{x: |x|<R\}, \qquad \omega_N R^N \leqq A,$$

$$\tilde{g}(x, 0:R)=0 \quad \text{on } \{x: |x|=R\}.$$

Let $\tilde{h}(x:R)=\tilde{g}(x, 0:R)-s(x, 0)$. If $\tilde{g}(x, 0:R)$ exists in the classical sense, then $\tilde{h}(x:R)$ is continuous in $\{x: |x|<R\}$ and $\lim_{x\to 0}\tilde{h}(x:R)=\tilde{h}(0:R)$. As in § 1.1 we define $h(x, y)$ to be the corresponding quantity for $g(x, y)$. We then have

LEMMA 2.3. *There exists a unique value R_y such that $\tilde{g}(x, 0:R_y)>0$ and such that $\tilde{h}(0:R_y)=h(y, y)$, provided $p(x)$ is of one sign.*

Proof. For R sufficiently small, $\tilde{g}(x, 0:R)$ is positive. This follows from Lemma 1.1 and the monotone dependence of the corresponding eigenvalue with respect to the domain. In the interval $(0, R_0)$ where $\tilde{g}(x, 0:R)$ is positive, $\tilde{h}(0:R)$ is a continuous, increasing function of R. We first show that $\tilde{h}(0:R)$ tends to $-\infty$ as R approaches zero. Let $p_0\geqq 0$ be an upper bound for $p^*(x)$, and consider the boundary value problem

$$\Delta h_0+p_0(h_0+s(x, 0))=0 \quad \text{in } \{|x|<\varepsilon\}, \qquad h_0=-s(x, 0) \quad \text{on } \{|x|=\varepsilon\}.$$

We choose ε small enough to obtain a positive Green's function $g(x, y:p_0)$. Because of (B) in § 1.2 we have $\tilde{h}(0:\varepsilon)\leqq h_0(0)$. Let $m=\sup h_0$, and $\mathcal{G}(x, 0)$ to be the Green's function of the Laplace operator in $\{|x|<\varepsilon\}$. In view of the symmetry of the problem, $m=h_0(0)$. Hence,

$$m=\int_{\{|x|<\varepsilon\}}\mathcal{G}(x, 0)(-\Delta h_0)\,dx+\oint_{\{|x|=\varepsilon\}}\frac{\partial\mathcal{G}}{\partial n}s(x, 0)\,ds\leqq\oint_{\{|x|<\varepsilon\}}\mathcal{G}(x, 0)(m+s)p_0\,dx-s(\varepsilon, 0)$$

$$=\frac{\varepsilon^2 p_0}{2N}m+c(\varepsilon)-s(\varepsilon, 0).$$

An explicit computation shows that $\lim_{\varepsilon\to 0}c(\varepsilon)=0$. Thus, $m\to-\infty$ if $\varepsilon\to 0$. For ε sufficiently small, $m<h(y, y)$, and $\delta(\varepsilon)=h(y, y)-\tilde{h}(0:\varepsilon)>0$. In order to show that $\delta(R)$ takes also negative values, we distinguish two cases.

(i) Suppose that $\tilde{g}(x, 0:R)$ exists and is positive in the whole interval $(0, (A/\omega_N)^{1/N})$. Observing that $\lim_{a\to 0}[H'(a)-H^{*\prime}(a)]=\rho(0)\delta((A/\omega_N)^{1/N})$ we conclude directly from Theorem 2.1 or (2.8) that $\delta((A/\omega_N)^{1/N})\leqq 0$.

(ii) If $\tilde{g}(x, 0:R)$ is not defined and positive in the whole interval $(0, (A/\omega_N)^{1/N})$, then there exists a value R_0 such that $\lim_{R\to R_0}\tilde{h}(0:R)=+\infty$. In both cases $\delta(R)$ is nonpositive for some value of R. Because of its monotonicity $\delta(R)$ vanishes exactly once.

In the sequel we write $\tilde{g}(x, 0)$ for $\tilde{g}(x, 0:R_y)$. Lemma 2.2 leads to the following estimate.

THEOREM 2.2. *If $p(x)$ is of one sign, then for fixed $y\in D$*

$$0\leqq\tilde{g}(x, 0)\leqq g^*(x, y) \quad \text{in } D_y=\{x: |x|<R_y\}.$$

Proof. Let $\tilde{H}(a)$ and $\tilde{\mu}(a)$ be the functions corresponding to $H(a)$ and $\mu(a)$ with $g(x, y)$ replaced by $\tilde{g}(x, 0)$. Proceeding the same way as in the proof of Lemma 2.2 we find that

$$\tilde{F}'(a)=\left[\frac{\tilde{H}H'}{\rho}-\frac{\tilde{H}'H}{\rho}\right]'\geqq 0 \quad \text{provided} \quad \rho(a)\neq 0.$$

From the definition of R_y it follows that

$$\lim_{a\to 0}\tilde{F}(a)=\lim_{a\to 0}\left\{\frac{\tilde{H}(a)H'(a)}{\rho(a)}-\frac{H(a)\tilde{H}'(a)}{\rho(a)}\right\}=0.$$

Let us first exclude the case where $\rho(a)=0$ for some a. We deduce from $\tilde{F}(a)\geqq 0$, that $H(a)/\tilde{H}(a)\leqq \mu(a)/\tilde{\mu}(a)$.

(i) Let $\rho(a)>0$. Then $\tilde{F}(a)\rho(a)/\tilde{H}(a)^2=[H(a)/\tilde{H}(a)]'\geqq 0$ and $H(a)\geqq \tilde{H}(a)$. This estimate together with the previous observation yields $\mu(a)\geqq\tilde{\mu}(a)$. If $\rho(a)\geqq 0$ the arguments are very similar, we have just to use (2.3) in the interval where $\rho(a)$ vanishes.

(ii) If $\rho(a)<0$, we conclude that $H(a)\leqq \tilde{H}(a)$. The difference $H-\tilde{H}=D(a)$ satisfies

$$\left(\frac{D'}{\rho}\right)'+q^{-1}(a)D\geqq 0, \qquad D'(0)=0 \quad \text{and} \quad D(a)\leqq 0.$$

Hence,

$$\mu(a)-\tilde{\mu}(a)\geqq -\int_0^a q^{-1}(b)D(b)\,db\geqq 0.$$

The case $p(x)\leqq 0$ can be treated similarly.

3. Applications.

3.1. Estimates for integrals involving the Green's functions. Let us use the same notations as before. Note that $g(x, y)$ and $g^*(x, y)$ are equimeasurable and that

$$\int_D \Phi(g(x, y))\,dx=\int_{D^*}\Phi(g^*(x, y))\,dx \tag{3.1}$$

for every function $\Phi(t)$ for which the integrals in (3.1) make sense. (3.1) together with Theorems 2.1 and 2.2 yields

THEOREM 3.1. *If $p(x)\geqq 0$ and $G(x, 0)\geqq 0$, then we have for every nondecreasing function $\Phi(t)$*

$$\int_{D_y}\Phi[\tilde{g}(x, 0)]\,dx\leqq\int_D\Phi[g(x, y)]\,dx\leqq\int_{D^*}\Phi[G(x, 0)]\,dx.$$

Remark. For $p\equiv 0$ and $\Phi(t)=t^k$, the upper bound was already established by Weinberger [16].

For negative $p(x)$ only a weaker result holds, namely

THEOREM 3.2. *If $p(x)\leqq 0$, then for every nondecreasing function $\Phi(t)$*

$$\int_{D_y}\Phi[\tilde{g}(x, 0)]\,dx\leqq\int_D\Phi[g(x, y)]\,dx.$$

This statement is an immediate consequence of (3.1) and Theorem 2.2. An interesting upper bound for integrals of the type considered previously, can only be given in the special case $p(x) = \text{const}$. We note the following corollary which may be deduced from (2.8).

COROLLARY 3.1. *Let $\Phi(t)$ be a nondecreasing convex function, for which all integrals used in the proof are well-defined, and let $p(x) = \alpha \in \mathbb{R}$ and $G(x, 0) \geqq 0$, then*

$$\int_D \Phi[g(x, y)]\, dx \leqq \int_{D^*} \Phi[G(x, 0)]\, dx.$$

Proof. By the convexity of $\Phi(t)$,

$$\begin{aligned}\int_{D^*} &\{\Phi[G(x, 0)] - \Phi[g^*(x, y)]\}\, dx \\ &\geqq \int_{D^*} \Phi'(g^*(x, 0))[G - g^*]\, dx \\ &= \Phi'(\mu(a))\left[\frac{H^*(a)}{\alpha} - \frac{H(a)}{\alpha}\right]_0^A - \frac{1}{\alpha}\int_0^A \Phi''(\mu)[H^*(a) - H(a)]\frac{d\mu}{da}\, da.\end{aligned}$$

Because of (2.8) and because $(d\mu/da) \leqq 0$, the whole expression is nonnegative, provided $\lim_{a\to 0} \Phi'(\mu(a))\{H^*(a) - H(a)\} = 0$.

3.2. A result on D_y. The aim of this section is to study R_y for the special case where $p = \alpha$, α being an arbitrary real number. We denote by $g_\alpha(x, y)$ the corresponding Green's function in D, and by $\tilde{g}_\alpha(x, 0)$ the solution of $\Delta\tilde{g}_\alpha + \alpha\tilde{g}_\alpha = -\delta_0$ in D_y, $\tilde{g}_\alpha = 0$ on D_y. Clearly, D_y depends on α $[D_y = D_y(\alpha)]$.

LEMMA 3.1. *Let α vary in the range of values $(-\infty, \lambda)$ for which a positive Green's function $g_\alpha(x, y)$ exists. Then $D_y(\alpha)$ is monotonically increasing in $(-\infty, \lambda)$.*

Proof. Let $\alpha_0 \in (-\infty, \lambda)$ be any fixed number and put $G_0(x, 0:\alpha)$ for the positive solution of $\Delta v + \alpha v = -\delta_0$ in $D_y(\alpha_0)$, $v = 0$ on $\partial D_y(\alpha_0)$. In addition, let $d(\alpha) = g_\alpha(y, y) - G_0(0, 0:\alpha)$. Since the singularities of g and G_0 are of the same type, $d(\alpha)$ is a continuous function with $d(\alpha_0) = 0$. By Lemma 1.2

$$\dot{d}(\alpha_0) = \int_D g_{\alpha_0}(y, z)^2\, dz - \int_{D_y(\alpha_0)} G_0(z, 0:\alpha_0)^2\, dz.$$

Applying the integral estimates of the preceding section with $\Phi(t) = t^2$, we deduce that $\dot{d}(\alpha_0) \geqq 0$. Since the Green's functions are increasing with the domain, we must have $D_y(\alpha) \supseteq D_y(\alpha_0)$ for any $\alpha \geqq \alpha_0$ and α close to α_0. Since the argument holds for all α_0, the assertion of Lemma 3.1 is established.

From this result we obtain immediately the

COROLLARY 3.2. *If $\alpha \geqq 0$, then $D_y(0) \subseteq D_y(\alpha)$, $D_y(0)$ is the domain related with the Green's function of the Laplace operator.*

This corollary is useful especially for $N = 2$ where $D_y(0)$ is known for many domains [14]. Explicit values for the corresponding quantity in three dimensions are much more difficult to obtain. Upper bounds for the radius of $D_y(0)$ are computed in [11]. Lower bounds which are needed in our context seem not to be available yet.

3.3. Estimates for solutions of boundary value problems. Consider the boundary value problems

$$\Delta u + p(x)u + q(x) = 0 \quad \text{in } D, \qquad u = 0 \quad \text{on } \partial D. \tag{3.2}$$

and

$$\Delta U + p^*(x)U + q^*(x) = 0 \quad \text{in } D^*, \qquad U = 0 \quad \text{on } \partial D^*. \tag{3.3}$$

$g(x, y)$ and $G(x, y)$ are the corresponding Green's functions.

THEOREM 3.3. *If $q(x) \geqq 0$, $p(x) \geqq 0$ and $G(x, y) \geqq 0$, then*

$$u(x) \leqq \sup_{x \in D^*} U(x).$$

Proof. Because of the integral representation and the rearrangement theorem [9]

$$u(x) = \int_D g(x, y) q(y)\, dy \leqq \int_{D^*} g^*(x, y) q^*(y)\, dy.$$

Here, $g^*(x, y)$ denotes the symmetrized function with respect to y. The assertion follows from the symmetry of the Green's functions and from Theorem 2.1.

This theorem generalizes a result (Theorem 1.1) of [1]. For $p \leqq 0$, we have the weaker result

THEOREM 3.4. *If $p(x) \leqq 0$ and $q(x) = 1$, then*

$$\sup_{x \in D} u(x) \leqq U(0).$$

The proof proceeds as the one of Theorem 3.3. We have only to use inequality (2.8) instead of Theorem 2.1.

A particular case of this theorem is found in [1]. Moreover we have

THEOREM 3.5. *If $q = 1$, $p(x)$ is of one sign and $g(x, y) \geqq 0$,*

$$u(y) \geqq U_y(0),$$

where U_y is the solution of (3.3) *with D^* replaced by D_y.*

Proof. By Theorem 3.2 we have

$$u(y) = \int_D g(x, y)\, dx \geqq \int_{D_y} \tilde{g}(x, 0)\, dx.$$

We are now in a position to give an answer to our original problem,

$$\Delta u + \alpha u + 1 = 0 \quad \text{in } D, \qquad u = 0 \quad \text{on } \partial D.$$

Suppose that $\alpha > 0$, and that the Green's function $g(x, y : \alpha)$ is positive. In view of the Rayleigh–Faber–Krahn inequality [14] and Lemma 1.1 this is certainly the case when $\alpha < [\omega_N/A]^{2/N} j^2_{(N-2)/2}$, j_k is the first zero of the Bessel function J_k. Let $\dot{R}$ be the maximal conformal or harmonic radius of D. Then by Lemma 3.1 and Theorem 3.5 we get

COROLLARY 3.3. *Under the conditions stated above we have*

$$\max u(x) \geqq \begin{cases} \dfrac{1}{\alpha}\left(\dfrac{\sqrt{\alpha}\dot{R}}{2J_0(\sqrt{\alpha}\dot{R})} - 1\right) & \text{if } N = 2, \\[2ex] \dfrac{1}{\alpha}\left(\dfrac{\sqrt{\alpha}\dot{R}}{\sin(\sqrt{\alpha}\dot{R})} - 1\right) & \text{if } N = 3. \end{cases}$$

Equality holds if D is a sphere.

Acknowledgment. The author would like to thank the referee for suggestions on improving the presentation.

REFERENCES

[1] C. BANDLE, *Bounds for the solutions of boundary value problems*, J. Math. Anal. Appl., 54 (1976), pp. 706–716.
[2] ———, *On symmetrizations in parabolic equations*, J. Analyse Math., 30 (1976), pp. 98–112.
[3] ———, *Isoperimetric inequalities for a nonlinear eigenvalue problem*, Proc. Amer. Math. Soc., 56 (1976), pp. 243–246.
[4] R. BELLMAN, *On the non-negativity of Green's functions*, Boll. Un. Mat. Ital., 12 (1957), pp. 411–413.
[5] S. BERGMAN AND M. SCHIFFER, *Kernel Functions and Elliptic Differential Equations in Mathematical Physics*, New York, 1953.
[6] R. COURANT AND D. HILBERT, *Methods of Mathematical Physics*, vol. I, New York, 1953.
[7] P. R. GARABEDIAN, *Partial Differential Equations*, New York, 1964.
[8] P. R. GARABEDIAN AND M. M. SCHIFFER, *Convexity of Domain Functional*, J. Analyse Math., 2 (1952–53), pp. 281–369.
[9] G. H. HARDY, J. E. LITTLEWOOD AND G. PÓLYA, *Inequalities*, Cambridge University Press, London, 1967.
[10] J. HERSCH, *Sur la fréquence fondamentale d'une membrane vibrante: évaluation par défaut et principe du maximum*, Z. Angew. Math. Phys., 11 (1960), pp. 387–413.
[11] ———, *Transplantation harmonique, transplantation par modules et théorèmes isopérimetriques*, Comment. Math. Helv., 44 (1969), pp. 354–366.
[12] J. M. LUTTINGER AND R. FRIEDBERG, *Some functional inequalities, with applications to Green's function*, Arch. Rational Mech. Anal., 61 (1976), pp. 187–195.
[13] L. E. PAYNE, *Some Isoperimetric Inequalities in the Torsion Problem for Multiply Connected Regions*, Studies in Mathematical Analysis and Related Topics, Stanford, CA, 1962.
[14] G. PÓLYA AND G. SZEGÖ, *Isoperimetric Inequalities in Mathematical Physics*, Princeton, 1951.
[15] B. SCHWARZ, *Bounds for the principal frequency of the nonhomogeneous membrane and for the generalized Dirichlet integral*, Pacific J. Math., 7 (1957), pp. 1653–1676.
[16] H. F. WEINBERGER, *Symmetrization in Uniformly Elliptic Problems*, Studies in Mathematical Analysis and Related Problems, Stanford, CA, 1962.

SIAM J. MATH. ANAL.
Vol. 9, No. 6, December 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0013 $01.00/0

AN N-DIMENSIONAL EXTENSION OF THE STURM SEPARATION AND COMPARISON THEORY TO A CLASS OF NONSELFADJOINT SYSTEMS*

SHAIR AHMAD† AND A. C. LAZER‡

Abstract. Sturmian theory is extended to nonselfadjoint second order linear homogeneous systems. Almost all the results obtained are new even in the selfadjoint case.

1. Introduction and summary. The differential equations to be considered in this paper have the form

$$y''(t)+A(t)y(t)=0,$$

where y is a real n-dimensional vector and $A(t)$ is a real $n\times n$ matrix continuous on some interval. It is further assumed that the off-diagonal elements of $A(t)$ are always nonnegative.

For the case $n=1$ this equation has been studied extensively beginning with a famous paper by Sturm [10] in 1836. More recently there have been various extensions of the Sturmian theory to selfadjoint systems of second order linear differential equations, initiated by M. Morse [6] in 1930. Further extensions were subsequently given by Birkhoff and Hestenes [3], Reid [8], and others. For accounts of this work we refer the reader to the books of Coppel [4], Morse [7] and Reid [9].

The selfadjoint systems of differential equations considered in the works we have cited generally have a more complex form than the type we consider but include this type only when the matrix $A(t)$ is symmetric. The extensions of the Sturmian theory to selfadjoint systems are consequences of the fact that the selfadjoint systems are the Euler–Lagrange equations of certain quadratic functionals. The variational principles from which these extensions have been derived seem to be of no value if $A(t)$ is nonsymmetric.

In this paper we give several theorems pertaining to the system given above which are either equivalent to, or are implied by the Sturm separation and comparison theorems when $n=1$. Most of these results appear to be new even in the selfadjoint case. Namely, we prove:

THEOREM 1. *Let $A(t)$ and $B(t)$ be continuous $n\times n$ matrices defined on $[a, b]$ such that if $A(t)=(a_{ij}(t))$ and $B(t)=(b_{ij}(t))$ then $a_{ij}(t)\geqq b_{ij}(t)$ for $1\leqq i, j\leqq n$, $t\in[a, b]$ and such that $b_{ij}(t)\geqq 0$ for $i\neq j$ and $t\in[a, b]$. Assume that for some $\bar{t}\in[a, b]$, $a_{ij}(\bar{t})>b_{ij}(\bar{t})$, $1\leqq i, j\leqq n$. If there exists a nontrivial solution of*

$$x''(t)+B(t)x(t)=0$$

such that $x(a)=x(b)=0$, then there exists a nontrivial solution of

$$y''(t)+A(t)y(t)=0$$

such that $y(a)=y(c)=0$ with $a<c<b$.

If $n=1$ this result is equivalent to the Sturm comparison theorem. Recalling that when $n=1$ any two solutions with a common zero are linearly dependent, we see that the next two results generalize the Sturm separation theorem.

* Received by the editors September 20, 1976, and in revised form March 11, 1977.
† Department of Mathematics, Oklahoma State University, Stillwater, Oklahoma 74074.
‡ Department of Mathematics, University of Cincinnati, Cincinnati, Ohio 45221.

THEOREM 2. *Assume that the* $n \times n$ *matrix* $B(t) = (b_{ij}(t))$ *is continuous on* $[a, b]$ *and that* $b_{ij}(t) \geqq 0$ *if* $i \neq j$, $1 \leqq i, j \leqq n$. *Assume that there exists a nontrivial solution of*

$$x''(t) + B(t)x(t) = 0 \tag{S}$$

with $x(a) = x(b) = 0$ *and there exists no nontrivial solution* $y(t)$ *with* $y(a) = y(c) = 0$ *if* $a < c < b$. *If* $a \leqq t_1 < t_2 < b$ *then there exists no nontrivial solution* $y(t)$ *of* (S) *with* $y(t_1) = y(t_2) = 0$.

We recall (see [4] and [9]) that b is the first conjugate point of a relative to (S) if there is a nontrivial solution of (S) which vanishes at a and at b, and there is no nontrivial solution which vanishes at a and c for $a < c < b$. The equation (S) is said to be disconjugate on an interval I if no nontrivial solution vanishes more than once on I. For the selfadjoint case it is well-known that if b is the first conjugate point of a, then (S) is disconjugate on $[a, b)$. Theorem 2 shows this to be true also for the nonselfadjoint case.

THEOREM 3. *Let* $B(t)$ *satisfy the same conditions as in Theorem* 2. *Assume that there exists a nontrivial solution* $x(t)$ *of* (S) *with* $x(a) = x(b) = 0$. *If* $y(t) = \text{col}\,(y_1(t), \cdots, y_n(t))$ *is any solution of* (S), *one of the following must hold*:

(i) *There exist* k *and* $\bar{t} \in [a, b]$ *with* $y_k(\bar{t}) = 0$.

(ii) *There exist* k *and* l *with* $k < l$ *such that* $y_k(t)y_l(t) < 0$ *for all* $t \in [a, b]$.

We give a simple example related to the last theorem. Let

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

If $x(t) = \text{col}\,(\sin \sqrt{2}\, t, \sin \sqrt{2}\, t)$ then $x''(t) + Bx(t) = 0$ and $x(0) = x(\pi/\sqrt{2}) = 0$. If $y(t) = \text{col}\,(t, -t)$ then $y''(t) + B(t)y(t) = 0$, $y(0) = 0$, and $y(t) \neq 0$ for $t > 0$. Thus, for $n > 1$ it is possible, in Theorem 2, that the first condition be satisfied for a solution $y(t)$ independent of $x(t)$ and that no component of $y(t)$ vanish on the open interval (a, b). To illustrate the second alternative for the same example we observe that if $c \notin [0, \pi/\sqrt{2}]$ and $z(t) = \text{col}\,(t - c, c - t) = \text{col}\,(z_1(t), z_2(t))$ then $z''(t) + Bz(t) = 0$ and $z_1(t)z_2(t) < 0$ for $t \in [0, \pi/\sqrt{2}]$.

As a by-product of the methods used to establish the above theorems we prove an intermediate result which is a triviality when $n = 1$ but seems interesting when $n > 1$.

THEOREM 4. *Assume the hypotheses of Theorem* 2, *and let* b *be the first conjugate point of* a. *There exists a nontrivial solution* $u(t) = \text{col}\,(u_1(t), \cdots, u_n(t))$ *of* (S) *such that* $u(a) = u(b) = 0$ *and* $u_k(t) \geqq 0$, $k = 1, \cdots, n$ *and* $t \in [a, b]$.

This result was established by the authors in [1] for the case when $A(t)$ is a symmetric matrix, using methods from the calculus of variations.

Our principal tool used in deriving the above results is an extremal characterization of the smallest positive eigenvalue of a certain (nonselfadjoint) system of linear integral equations. We would like to acknowledge the influence of R. Bellman who gave an analogous extremal characterization of the Frobenius–Perron eigenvalue of a strictly positive matrix. (See [2, p. 288].) Although there is a definite connection between our preliminary lemmas and the theory of linear positive operators in partially ordered Banach spaces due to Krein and Rutman [5], our treatment will be simple and entirely self-contained.

2. Preliminary observations. Throughout this paper we shall make extensive use of Green's function for the boundary value problem $x''(t) = -f(t)$, $x(a) = x(b) = 0$

where $a < b$. Recall that

$$G(s,t)=\begin{cases}\dfrac{(s-a)(b-t)}{b-a}, & a \leqq s \leqq t \leqq b,\\[2mm] \dfrac{(t-a)(b-s)}{b-a}, & a \leqq t \leqq s \leqq b.\end{cases}$$

The function G is continuous on the square $a \leqq s \leqq b$, $a \leqq t \leqq b$, and for s and t in the same range

$$0 \leqq G(s,t) \leqq G(t,t) = \frac{(t-a)(b-t)}{(b-a)},$$

and hence

(1) $$G(s,t) \leqq \frac{b-a}{4}, \qquad a \leqq s,\ t \leqq b.$$

If $f(t)$ is a continuous real valued function defined for $a \leqq t \leqq b$ and if $x(t) = \int_a^b G(s,t) f(s)\,ds$ then, as is well-known (or by an easy calculation), $x(t)$ is of class C^2 on $[a,b]$, $x''(t) = -f(t)$ and $x(a) = x(b) = 0$. Moreover,

(2) $$x'(a) = \int_a^b \left(\frac{b-s}{b-a}\right) f(s)\,ds$$

and

(3) $$x'(b) = -\int_a^b \left(\frac{s-a}{b-a}\right) f(s)\,ds.$$

Conversely as there is only one solution of the boundary value problem $x''(t) = -f(t)$, $x(a) = x(b) = 0$, this solution must have the representation given above.

3. An extremal characterization of λ_0. If $x = \text{col}\,(x_1, \cdots, x_n) \in R^n$ and $y = \text{col}\,(y_1, \cdots, y_n) \in R^n$, we write $x \leqq y$ iff $x_k \leqq y_k$ for $k = 1, 2, \cdots, n$, and we write $x < y$ iff $x_k < y_k$ for $k = 1, 2, \cdots, n$. If $u\colon [a,b] \to R^n$ is continuous, we write $u \in K$ if $u(a) = u(b) = 0$ and $0 \leqq u(t)$ for all $t \in (a,b)$. Let $A(t) = (a_{ij}(t))$ be an $n \times n$ continuous matrix defined on $[a,b]$. Assume $a_{ij}(t) > 0$, $1 \leqq i, j \leqq n$ and $t \in [a,b]$, except possibly on a set of measure zero. If $u\colon [a,b] \to R^n$ is continuous, we define

(4) $$(Tu)(t) = \int_a^b G(s,t) A(s) u(s)\,ds.$$

It follows immediately that

(5) $$T(u+v) = Tu + Tv,$$

(6) $$T(cu) = cTu, \qquad c \in R,$$

(7) $$u \in K \Rightarrow Tu \in K,$$

(8) $$u \in K, \qquad u(t) \not\equiv 0 \Rightarrow 0 < (Tu)(t), \qquad t \in (a,b).$$

For $\lambda \in R$ we write $\lambda \in \Lambda$ if there exists $u \in K$, $u \not\equiv 0$, such that $u(t) \leqq \lambda (Tu)(t)$ for $t \in (a,b)$.

LEMMA 1. *$\Lambda \neq \varnothing$. If $\lambda_0 = \inf\{\lambda \mid \lambda \in \Lambda\}$, then $\lambda_0 > 0$.*

Proof. Let u be any nontrivial member of K such that $u \in C^1[a, b]$. From (2) and (4) we have

$$0 < \int_a^b \left(\frac{b-s}{b-a}\right) A(s)u(s)\, ds = (Tu)'(a).$$

Similarly, by (3)

$$(Tu)'(b) = -\int_a^b \left(\frac{s-a}{b-a}\right) A(s)u(s)\, ds < 0.$$

If $\lambda_1 > 0$ is sufficiently large then $u'(a) < \lambda_1 (Tu)'(a)$ and $\lambda_1 (Tu)'(b) < u'(b)$. As $u(a) = \lambda_1 (Tu)(a) = 0$ and $u(b) = \lambda_1 (Tu)(b) = 0$, there exists a number δ, $0 < \delta < (b-a)/2$, such that

$$u(t) < \lambda_1 (Tu)(t), \qquad t \in (a, a+\delta), \tag{9}$$

and

$$u(t) < \lambda_1 (Tu)(t), \qquad t \in (b-\delta, b). \tag{10}$$

If $t \in [a+\delta, b-\delta]$, it follows by (8) that $0 < (Tu)(t)$. Consequently if λ_2 is sufficiently large $u(t) < \lambda_2 (Tu)(t)$ for $t \in [a+\delta, b-\delta]$. Thus if $\lambda_3 = \max\{\lambda_1, \lambda_2\}$, $u(t) < \lambda_3 (Tu)(t)$ for all $t \in (a, b)$. Hence $\lambda_3 \in \Lambda$. To prove the second assertion, let $\lambda \in \Lambda$ and $u \in K$ such that $u(t) \not\equiv 0$ and

$$u(t) \leqq \lambda (Tu)(t) = \lambda \int_a^b G(s, t) A(s) u(s)\, ds, \qquad t \in [a, b]. \tag{11}$$

Let

$$\|A(t)\| = \max_{1 \leqq i \leqq n} \sum_{j=1}^n a_{ij}(t).$$

Let $u(t) = \operatorname{col}(u_1(t), \cdots, u_n(t))$. Let $1 \leqq k \leqq n$ and $\bar{t} \in [a, b]$ be such that $u_k(\bar{t}) = \max_{1 \leqq j \leqq n} \max_{t \in [a, b]} u_j(t)$. From (11) it follows that

$$\begin{aligned} u_k(\bar{t}) &\leqq \lambda \int_a^b G(s, \bar{t}) \sum_{j=1}^n a_{kj}(s) u_j(s)\, ds \\ &\leqq \lambda u_k(\bar{t}) \int_a^b G(s, \bar{t}) \sum_{j=1}^n a_{kj}(s)\, ds \\ &\leqq \lambda u_k(\bar{t}) \int_a^b G(s, \bar{t}) \|A(s)\|\, ds \\ &\leqq \lambda u_k(\bar{t}) \frac{(b-a)}{4} \int_a^b \|A(s)\|\, ds. \end{aligned}$$

Hence

$$\lambda \geqq \frac{4}{(b-a) \int_a^b \|A(s)\|\, ds},$$

and

$$\lambda_0 \geqq \frac{4}{(b-a) \int_a^b \|A(s)\|\, ds} > 0. \tag{12}$$

This estimate will be useful later.

LEMMA 2. *Let T and λ_0 be defined as above. If there exists $u \in K$ such that $u(t) \not\equiv 0$ and such that $u(t) \leqq \lambda_0(Tu)(t)$ for all $t \in (a, b)$ then $u(t) = \lambda_0(Tu)(t)$ for $t \in [a, b]$.*

Proof. Suppose, on the contrary, $\lambda_0(Tu)(t) - u(t) \not\equiv 0$. Let $w = Tu$. Since $\lambda_0 w - u \in K$ and $\lambda_0 w - u \not\equiv 0$, it follows from (8) that $0 < T(\lambda_0 w - u)(t)$ for $t \in (a, b)$. Thus by (5) and (6)

$$w(t) < \lambda_0(Tw)(t), \qquad t \in (a, b). \tag{13}$$

From (2) we have

$$\begin{aligned} w'(a) = (Tu)'(a) &= \int_a^b \left(\frac{b-s}{b-a}\right) A(s)u(s)\,ds \\ &< \lambda_0 \int_a^b \left(\frac{b-s}{b-a}\right) A(s)w(s)\,ds = \lambda_0(Tw)'(a). \end{aligned} \tag{14}$$

Similarly by (3),

$$\begin{aligned} \lambda_0(Tw)'(b) &= -\lambda_0 \int_a^b \left(\frac{s-a}{b-a}\right) A(s)w(s)\,ds \\ &< -\int_a^b \left(\frac{s-a}{b-a}\right) A(s)u(s)\,ds = (Tu)'(b) = w'(b). \end{aligned} \tag{15}$$

According to (14) and (15) there exists λ_1 with $0 < \lambda_1 < \lambda_0$ such that $w'(a) < \lambda_1(Tw)'(a)$ and $\lambda_1(Tw)'(b) < w'(b)$. As $w(a) = \lambda_1(Tw)(a) = w(b) = \lambda_1(Tw)(b) = 0$ there exists δ with $0 < \delta < (b-a)/2$ such that $w(t) < \lambda_1(Tw)(t)$ if $t \in (a, a+\delta]$ or $t \in [b-\delta, b)$. From (13) there exists λ_2, $0 < \lambda_2 < \lambda_0$, such that $w(t) < \lambda_2(Tw)(t)$ if $t \in [a+\delta, b-\delta]$. Thus if $\lambda_3 = \max\{\lambda_1, \lambda_2\}$, then $w(t) < \lambda_3(Tw)(t)$, $t \in (a, b)$, which contradicts the definition of λ_0. This contradiction proves that $u(t) = \lambda_0(Tu)(t)$ for $t \in [a, b]$.

LEMMA 3. *Let λ_0, T be as above. There exists $u \in K$, $u(t) \not\equiv 0$, such that $u(t) = \lambda_0(Tu)(t)$ for $t \in [a, b]$.*

Proof. Let $\{\lambda_m\}_1^\infty$ be a sequence in Λ and let $\{x_m\}_1^\infty$ be a sequence in K such that

$$x_m(t) \leqq \lambda_m(Tx_m)(t) \tag{16}$$

for $t \in [a, b]$ with $x_m(t) \not\equiv 0$, and

$$\lim_{m\to\infty} \lambda_m = \lambda_0. \tag{17}$$

By multiplying each x_m by a suitable positive constant we may assume that

$$\sum_{i=1}^n \sum_{j=1}^n \int_a^b a_{ij}(s)x_{mj}(s)\,ds = 1, \tag{18}$$

for each m, $m = 1, 2, \cdots$, where

$$x_m(t) = \operatorname{col}(x_{m1}(t), \cdots, x_{mn}(t)). \tag{19}$$

For each $m \geqq 1$ define

$$u_m(t) = (Tx_m)(t). \tag{20}$$

According to (16), $\lambda_m u_m - x_m \in K$. Hence, by (7), $T(\lambda_m u_m - x_m) = \lambda_m Tu_m - u_m \in K$. Hence, for $t \in [a, b]$ we have

$$u_m(t) \leqq \lambda_m(Tu_m)(t), \qquad m \geqq 1. \tag{21}$$

We claim that the elements of the vectors $\{u_m(t)\}_1^\infty$ are equicontinuous and uniformly bounded on $[a, b]$. To see this, let $u_m(t)=\mathrm{col}\,(u_{m1}(t), \cdots, u_{mn}(t))$. From (20)

$$0 \leqq u_{mk}(t)=\int_a^b G(s, t) \sum_{j=1}^n a_{kj}(s)x_j(s)\,ds. \tag{22}$$

Thus, by (1) and (18)

$$0 \leqq u_{mk}(t) \leqq \frac{b-a}{4}\int_a^b \sum_{j=1}^n a_{kj}(s)x_j(s)\,ds \leqq \frac{b-a}{4}$$

which shows that $\{u_{mk}(t)\}_{m=1}^\infty$ is a uniformly bounded sequence for $k=1, 2, \cdots, n$. Let $\varepsilon>0$. As G is uniformly continuous on $[a, b]\times[a, b]$ there exists $\delta>0$ such that if $t_1\in[a, b]$, $t_2\in[a, b]$ and $|t_1-t_2|<\delta$ then $|G(t_1, s)-G(t_2, s)|<\varepsilon$ for $s\in[a, b]$. Thus, if $|t_1-t_2|<\delta$, $m\geqq 1$, and $1\leqq k\leqq n$, from (22) we have

$$\begin{aligned}|u_{mk}(t_1)-u_{mk}(t_2)| &= \left|\int_a^b (G(s, t_1)-G(s, t_2)) \sum_{j=1}^n a_{kj}(s)x_j(s)\,ds\right| \\ &\leqq \int_a^b |G(s, t_1)-G(s, t_2)| \sum_{j=1}^n a_{kj}(s)x_j(s)\,ds \\ &< \varepsilon\int_a^b \sum_{j=1}^n a_{kj}(s)x_j(s)\,ds \leqq \varepsilon.\end{aligned}$$

By Ascoli's lemma we may assume without loss of generality that $\lim_{m\to\infty} u_m(t)=u(t)$ uniformly on $[a, b]$. Hence, according to (21)

$$u(t) \leqq \lim_{m\to\infty} \lambda_m(Tu_m)(t)=\lambda_0(Tu)(t), \qquad t\in[a, b]. \tag{23}$$

Suppose it were the case that $u(t)=0$ for all $t\in[a, b]$. From (16), $0\leqq x_m(t)\leqq \lambda_m u_m(t)$; hence $\lim_{m\to\infty} x_m(t)=0$ uniformly on $[a, b]$. Therefore,

$$\lim_{m\to\infty} \sum_{i=1}^n \sum_{j=1}^m \int_a^b a_{ij}(s)x_{mj}(s)\,ds=0,$$

contradicting (18). This proves that $u(t)\not\equiv 0$. Thus, by (23) and Lemma 2 it follows that $u(t)=\lambda_0(Tu)(t)$, $t\in[a, b]$ and the result is established.

LEMMA 4. *If there exists* $\lambda_1\in\Lambda$ *and* $w\in K$, $w(t)\not\equiv 0$, *such that*

$$w(t)=\lambda_1(Tw)(t) \quad \text{for } t\in[a, b], \tag{24}$$

then $\lambda_1=\lambda_0$.

Proof. Since $\lambda_1\in\Lambda$, $\lambda_1\geqq\lambda_0$. Suppose, contrary to the claim, $\lambda_1>\lambda_0$. By Lemma 3 there exists $u\in K$, $u\not\equiv 0$, such that $u(t)=\lambda_0(Tu)(t)$ for $t\in[a, b]$. Since according to (8), $(Tu)(t)>0$ for $t\in(a, b)$ we see that

$$0<u(t)<\lambda_1(Tu)(t), \qquad t\in(a, b). \tag{25}$$

Moreover, by (2),

$$0<u'(a)=\lambda_0\int_a^b \left(\frac{b-s}{b-a}\right)A(s)u(s)\,ds<\lambda_1\int_a^b \left(\frac{b-s}{b-a}\right)A(s)u(s)\,ds. \tag{26}$$

Similarly, using (3), it follows that

$$-\lambda_1\int_a^b \left(\frac{s-a}{b-a}\right)A(s)u(s)\,ds<u'(b)<0. \tag{27}$$

Similar considerations show that

(28) $$0 < w'(a), \quad w'(b) < 0, \quad w(t) > 0, \qquad t \in (a, b).$$

As $u(a) = w(a) = u(b) = w(b) = 0$, it follows from (28) that if $\alpha > 0$ is sufficiently small then

(29) $$0 < w'(a) - \alpha u'(a), \qquad w'(b) - \alpha u'(b) < 0,$$

and

(30) $$0 < w(t) - \alpha u(t), \qquad t \in (a, b).$$

If $\bar{\alpha} > 0$ is the least upper bound of the numbers α such that (29) and (30) hold, then by continuity

(31) $$0 \leqq w'(a) - \bar{\alpha} u'(a), \qquad w'(b) - \bar{\alpha} u'(b) \leqq 0,$$

and

(32) $$0 \leqq w(t) - \bar{\alpha} u(t), \qquad t \in (a, b).$$

Furthermore, at least one of the following possibilities must occur: For some k with $1 \leqq k \leqq n$ either

(33a) $$w_k'(a) - \bar{\alpha} u_k'(a) = 0,$$

(33b) $$w_k'(b) - \bar{\alpha} u_k'(b) = 0,$$

or

(33c) $$w_k(\bar{t}) - \bar{\alpha} u_k(\bar{t}) = 0,$$

for some $\bar{t} \in (a, b)$, where $u = \operatorname{col}(u_1, \cdots, u_n)$, $w = \operatorname{col}(w_1, \cdots, w_n)$. Otherwise we could find $\alpha > \bar{\alpha}$ such that (31) and (32) hold. We now show that all three possibilities are incompatible with previous inequalities. Since $\bar{\alpha} > 0$ it follows from (26) and (32) that

$$\begin{aligned} w'(a) - \bar{\alpha} u'(a) &= \lambda_1 (Tw)'(a) - \bar{\alpha} u'(a) \\ &= \lambda_1 \int_a^b \left(\frac{b-s}{b-a}\right) A(s) w(s)\, ds - \bar{\alpha} u'(a) \\ &> \lambda_1 \int_a^b \left(\frac{b-s}{b-a}\right) A(s)[w(s) - \bar{\alpha} u(s)]\, ds \geqq 0. \end{aligned}$$

Therefore, (33a) is impossible. Similarly, from (27) and (32),

$$\begin{aligned} w'(b) - \bar{\alpha} u'(b) &= -\lambda_1 \int_a^b \left(\frac{s-a}{b-a}\right) A(s) w(s)\, ds - \bar{\alpha} u'(b) \\ &< -\lambda_1 \int_a^b \left(\frac{s-a}{s-b}\right) A(s)[w(s) - \bar{\alpha} u(s)]\, ds \leqq 0. \end{aligned}$$

Consequently, (33b) is impossible. Finally if $\bar{t} \in (a, b)$ it follows from (24), (25) and (32) that

$$w(\bar{t}) - \bar{\alpha} u(\bar{t}) > \lambda_1 \int_a^b G(s, \bar{t}) A(s)[w(s) - \bar{\alpha} u(s)]\, ds \geqq 0,$$

which rules out (33c). This contradiction shows that $\lambda_1 = \lambda_0$. A slight modification of the proof shows that $w(t)$ is a multiple of $u(t)$ but this result will not be needed.

4. Monotonicity of λ_0. In this section we again assume that a and b are two numbers with $a<b$. However, we let b vary. Accordingly, we let $G(s,t,b)$ denote the Green's function for the interval $[a,b]$. The matrix $A(t)=(a_{ij}(t))$ is assumed to be continuous on $[a,\infty)$ with $a_{ij}(t)>0$, $1\leqq i,j\leqq n$ except at isolated points. The sets $\Lambda(b)$, $K(b)$, and the number $\lambda_0(b)$ depending on b are defined as before.

Our next lemma follows trivially.

LEMMA 5. *If $a<b_1<b_2$ and $t\in(a,b_1)$ then $G(t,s,b_1)<G(t,s,b_2)$ for $s\in(a,b_1]$.*

LEMMA 6. *If $a<b_1<b_2$ then $\lambda_0(b_2)<\lambda_0(b_1)$.*

Proof. According to Lemma 3 there exists $u\in K(b_1)$ such that $u(t)\not\equiv 0$ on $[a,b_1]$ and such that

$$u(t)=\lambda_0(b_1)\int_a^{b_1} G(s,t,b_1)A(s)u(s)\,ds.$$

Define $\hat{u}\in K(b_2)$ as follows:

$$\hat{u}(t)=\begin{cases} u(t), & a\leqq t<b_1,\\ 0, & b_1\leqq t\leqq b_2.\end{cases}$$

If $a<t\leqq b_1$ then by Lemma 5

$$\begin{aligned}\hat{u}(t)=u(t)&=\lambda_0(b_1)\int_a^{b_1} G(s,t,b_1)A(s)u(s)\,ds\\ &<\lambda_0(b_1)\int_a^{b_1} G(s,t,b_2)A(s)\hat{u}(s)\,ds\\ &=\lambda_0(b_1)\int_a^{b_2} G(s,t,b_2)A(s)\hat{u}(s)\,ds.\end{aligned}$$

If $b_1\leqq t<b_2$ then

$$\hat{u}(t)=0<\lambda_0(b_1)\int_a^{b_2} G(s,t,b_2)A(s)\hat{u}(s)\,ds.$$

Hence, $\lambda_0(b_2)\leqq\lambda_0(b_1)$, by definition. The assumption of equality gives $\hat{u}\leqq\lambda_0(b_2)T(\hat{u})$, where T refers to $[a,b_2]$, and Lemma 2 gives $\hat{u}=\lambda_0(b_2)T(\hat{u})$, contrary to $T(\hat{u})(b_1)>0$ which was shown above.

LEMMA 7. *The function $\lambda_0(b)$ is continuous on (a,∞) and $\lambda_0(b)\to\infty$ as $b\to a$.*

Proof. From the estimate

$$\lambda_0(b)\geqq\frac{4}{(b-a)\int_a^b\|A(s)\|\,ds}, \tag{12}$$

we see that $\lambda_0(b)\to\infty$ as $b\to a$. To establish continuity of $\lambda_0(b)$, fix a number $\bar{b}>a$. We shall show that $\lambda_0(b)$ is continuous from both the left and the right at $b=\bar{b}$. Since $\lambda_0(b)$ is nonincreasing on (a,∞) it follows that $\lim_{b\to\bar{b}+0}\lambda_0(b)=\lambda_1$ exists and $\lambda_1\leqq\lambda_0(\bar{b})$. Let $\{b_m\}_1^\infty$ be a sequence with $\bar{b}<b_{m+1}<b_m$ and $\lim_{m\to\infty}b_m=\bar{b}$. According to Lemma 3, for each $m\geqq 1$ there exists $u_m\in K(b_m)$ with $u_m\not\equiv 0$ such that

$$u_m(t)=\lambda_0(b_m)\int_a^{b_m} G(s,t,b_m)A(s)u_m(s)\,ds.$$

Hence, for $t\in[a,b_m]$, $u_m''(t)+\lambda_0(b_m)A(t)u_m(t)=0$ and $u_m(a)=u_m(b_m)=0$. By the uniqueness theorem $u_m'(a)\neq 0$, so by multiplying u_m by a suitable positive constant we may assume without loss of generality that $\|u_m'(a)\|=1$, where $\|\cdot\|$ denotes the usual

Euclidean norm. By choosing a suitable subsequence of the sequence $\{u_m(t)\}_1^\infty$ we may assume, without loss of generality, that $\lim_{m\to\infty} u_m'(a)=c\in R^n$ with $\|c\|=1$. If $w(t)$ denotes the solution of the initial value problem

$$\begin{aligned} &w''+\lambda_1 A(t)w=0,\\ &w(a)=0, \qquad w'(a)=c\neq 0, \end{aligned} \tag{34}$$

then by a standard result concerning continuity of solutions of differential equations with respect to initial conditions and with respect to parameters (see for example [9]) it follows that $\lim_{m\to\infty} u_m(t)=w(t)$ uniformly on compact subintervals of $[a,\infty)$. In particular since $u_m(t)\geqq 0$ for $a\leqq t\leqq b_m$ it follows that $w(t)\geqq 0$ on $[a,b]$ and $0=\lim_{m\to\infty} u_m(b_m)=w(\bar{b})$. Thus $w\in K(\bar{b})$, $w\not\equiv 0$ and according to (34)

$$w(t)=\lambda_1\int_a^b G(s,t)A(s)w(s)\,ds.$$

Thus, by Lemma 4, $\lambda_1=\lambda_0(\bar{b})$. This proves right-hand continuity of $\lambda_0(b)$ at $\bar{b}$. To establish left-hand continuity at $\bar{b}$ we observe that since $\lambda_0(b)$ is nonincreasing, $\lambda_2=\lim_{b\to\bar{b}-0}\lambda_0(b)$ exists, and a repetition of the previous argument shows that $\lambda_2=\lambda_0(\bar{b})$. This proves the result.

LEMMA 8. *Let $A(t)=(a_{ij}(t))$ and $\hat{A}(t)=(\hat{a}_{ij}(t))$ be $n\times n$ matrices which are continuous on $[a,b]$ and for $1\leqq i\leqq n$, $1\leqq j\leqq n$, $0<a_{ij}(t)\leqq \hat{a}_{ij}(t)$ on (a,b). For $u\in K(b)$ let*

$$(Tu)(t)=\int_a^b G(s,t)A(s)u(s)\,ds,$$

$$(\hat{T}u)(t)=\int_a^b G(s,t)\hat{A}(s)u(s)\,ds.$$

Let Λ be the set of numbers λ such that $u(t)\leqq\lambda(Tu)(t)$, $t\in(a,b)$, for some $u\in K(b)$, $u\not\equiv 0$, and let $\hat{\Lambda}$ be the set of numbers λ such that there exists $u\in K(b)$ such that $u(t)\leqq\lambda(\hat{T}u)(t)$, $t\in(a,b)$, for some $u\in K(b)$, $u\not\equiv 0$. If $\lambda_0(b)=\inf\{\lambda\,|\,\lambda\in\Lambda\}$ and $\hat{\lambda}_0(b)=\inf\{\lambda\,|\,\lambda\in\hat{\Lambda}\}$, then $\hat{\lambda}_0(b)\leqq\lambda_0(b)$.

Proof. According to Lemma 3 there exists $u\in K(b)$ such that $u=\lambda_0(b)Tu$, $u\not\equiv 0$. Hence for $t\in(a,b)$

$$u(t)=\lambda_0(b)\int_a^b G(s,t)A(s)u(s)\,ds\leqq\lambda_0(b)\int_a^b G(s,t)\hat{A}(s)u(s)\,ds.$$

Hence, $\lambda_0(b)\in\hat{\Lambda}$ and $\hat{\lambda}_0(b)=\inf\{\lambda\,|\,\lambda\in\hat{\Lambda}\}\leqq\lambda_0(b)$.

LEMMA 9. *Let $A(t)=(a_{ij}(t))$, $B(t)=(b_{ij}(t))$ be two continuous $n\times n$ matrices defined on $[a,b]$ such that $0\leqq b_{ij}(t)\leqq a_{ij}(t)$, $t\in[a,b]$, $1\leqq i\leqq n$, $1\leqq j\leqq n$ and for some $\bar{t}\in(a,b)$, $0\leqq b_{ij}(\bar{t})<a_{ij}(\bar{t})$, $1\leqq i\leqq n$, $1\leqq j\leqq n$. Suppose $x''+B(t)x=0$, $x(t)\not\equiv 0$, $x(a)=x(b)=0$.*

ASSERTION. *There exists a solution of $u''+A(t)u=0$, $u(a)=u(c)=0$, $u(t)\not\equiv 0$ with $a<c<b$, and $u\in K(c)$.*

Proof. We have for $t\in[a,b]$,

$$x(t)=\int_a^b G(s,t)B(s)x(s)\,ds.$$

If $x(t)=\operatorname{col}(x_1(t),\cdots,x_n(t))$, let $w(t)=\operatorname{col}(|x_1(t)|,\cdots,|x_n(t)|)$. Then $w\in K(b)$ and $w\not\equiv 0$.

For $k=1,\cdots,n$,

$$w_k(t)=|x_k(t)|=\left|\int_a^b G(s,t)\sum_{j=1}^n b_{kj}(s)x_j(s)\,ds\right|$$
$$\leqq\int_a^b G(s,t)\sum_{j=1}^n b_{kj}(s)|x_j(s)|\,ds$$
$$=\int_a^b G(s,t)\sum_{j=1}^n b_{kj}(s)w_j(s)\,ds.$$

Now by the uniqueness theorem for differential equations, the components of $w(t)$ cannot vanish simultaneously on any subinterval of $[a,b]$ since $x(t)\not\equiv 0$. Thus, since $b_{kj}(s)\leqq a_{kj}(s)$, $s\in(a,b)$, and $b_{kj}(\bar t)<a_{kj}(\bar t)$, we have

$$\int_a^b G(s,t)\sum_{j=1}^n b_{kj}(s)w_j(s)\,ds<\int_a^b G(s,t)\sum_{j=1}^n a_{kj}(s)w_j(s)\,ds$$

for $t\in(a,b)$. Hence, we have

$$0\leqq w(t)<\int_a^b G(s,t)A(s)w(s)\,ds \tag{35}$$

for $t\in(a,b)$

Since the elements of $A(t)$ are not strictly positive on $[a,b]$ we cannot use our previous results directly. For each integer $m=1,2,\cdots$, let $A_m(t)=(a_{ij}(t)+1/m)$. As the elements of A_m are strictly positive on $[a,b]$, our previous results are applicable. Clearly, for $m\geqq 1$,

$$0\leqq w(t)<\int_a^b G(s,t)A_m(s)w(s)\,ds, \tag{36}$$

for $t\in(a,b)$. For each $m\geqq 1$ and $d\in(a,b]$, define

$$(T_m^d u)(t)=\int_a^d G(s,t,d)A_m(s)u(s)\,ds$$

for $u\in K(d)$; let $\Lambda_m(d)$ be the set of numbers λ such that $u(t)\leqq\lambda(T_m^d u)(t)$ for $t\in[a,b]$, and let $\lambda_{0m}(d)=\inf\{\lambda\,|\,\lambda\in\Lambda_m(d)\}$. If $m_1<m_2$ then each element of $A_{m_1}(t)$ is greater than the corresponding element of $A_{m_2}(t)$, so by Lemma 8

$$m_1<m_2\Rightarrow\lambda_{0m_1}(d)\geqq\lambda_{0m_2}(d). \tag{37}$$

From (36) we see that $1\in\Lambda_m(b)$ for all m, and hence, $\lambda_{0m}(b)\leqq 1$ for all m. As $\lambda_{0m}(d)$ is continuous, decreasing in d, and $\lambda_{0m}(d)\to+\infty$ as $d\to a$, there exists a unique $d_m\in(a,b]$ such that $\lambda_{0m}(d_m)=1$. Moreover by (37) it follows that

$$a<d_{m_1}\leqq d_{m_2}\quad\text{if } m_1<m_2. \tag{38}$$

Hence, $\lim_{m\to\infty}d_m=c$ for some $c\in(a,b]$. By Lemma 3 there exists $u_m\in K(d_m)$, $u_m\not\equiv 0$, such that

$$u_m(t)=\lambda_{0m}(d_m)\int_a^{d_m}G(s,t,d_m)A_m(s)u_m(s)\,ds$$
$$=\int_a^{d_m}G(s,t,d_m)A_m(s)u_m(s)\,ds.$$

Hence $u_m''+A_mu_m=0$, $u_m(a)=u_m(d_m)=0$. Without loss of generality, as in the proof of Lemma 7, we may assume that $\lim_{m\to\infty} u_m'(a)=k\neq 0$. As $A_m(t)\to A(t)$ uniformly on $[a,\infty)$ it follows that if $u(t)$ is the solution of the initial value problem $u''+A(t)u=0$, $u(a)=0$, $u'(a)=k$ then $u_m(t)\to u(t)$ uniformly on compact subintervals of $[a,\infty)$. Hence, $u(c)=\lim_{m\to\infty} u_m(d_m)=0$; obviously $u\in K(c)$. To complete the proof, we must show that $c<b$. Assume on the contrary that $c=b$, so that

$$u(t)=\int_a^b G(s,t)A(s)u(s)\,ds. \tag{39}$$

Let

$$v(t)=\int_a^b G(s,t)A(s)w(s)\,ds. \tag{40}$$

Then v is of class C^2 on $[a,b]$. According to (35), $0\leqq w(t)<v(t)$, $r\in(a,b)$. Hence, by the nonnegativity of the elements of $A(s)$, $s\in(a,b)$, the strict positivity of the elements of $A(\bar{t})$, and the strict positivity of $G(s,t)$ for $a<s<b$, $a<t<b$, it follows that for $t\in(a,b)$,

$$v(t)=\int_a^b G(s,t)A(s)w(s)\,ds<\int_a^b G(s,t)A(s)v(s)\,ds. \tag{41}$$

Similarly,

$$v'(a)=\int_a^b \left(\frac{b-s}{b-a}\right)A(s)w(s)\,ds<\int_a^b \left(\frac{b-s}{b-a}\right)A(s)v(s)\,ds, \tag{42}$$

and

$$-\int_a^b \left(\frac{s-a}{b-a}\right)A(s)v(s)\,ds<-\int_a^b \left(\frac{s-a}{b-a}\right)A(s)w(s)\,ds=v'(b). \tag{43}$$

Since, by the uniqueness theorem, the components of $u(t)$ cannot vanish simultaneously on any open subinterval of (a,b), the same type of reasoning shows that

$$0<u(t),\qquad t\in(a,b), \tag{44}$$

$$0<\int_a^b \left(\frac{b-s}{b-a}\right)A(s)u(s)\,ds=u'(a), \tag{45}$$

$$u'(b)=-\int_a^b \left(\frac{s-a}{b-a}\right)A(s)u(s)\,ds<0. \tag{46}$$

Using (44), (45) and (46) and the exact same reasoning as in the proof of Lemma 4 we infer the existence of a number $\bar{\alpha}>0$ such that

$$0\leqq u(t)-\bar{\alpha}v(t),\qquad t\in(a,b), \tag{47}$$

$$0\leqq u'(a)-\bar{\alpha}v'(a),\qquad u'(b)-\bar{\alpha}v'(b)\leqq 0, \tag{48}$$

and such that for some k, $1\leqq k\leqq n$, one of the following three possibilities must hold: If $u=\operatorname{col}(u_1,\cdots,u_n)$, $v=\operatorname{col}(v_1,\cdots,v_n)$, either

$$u_k(\bar{t})-\bar{\alpha}v_k(\bar{t})=0\quad\text{for some }\bar{t},\ a<\bar{t}<b, \tag{49a}$$

$$u_k'(a)-\bar{\alpha}v_k'(a)=0, \tag{49b}$$

or

$$u_k'(b)-\bar{\alpha}v_k'(b)=0. \tag{49c}$$

However, as $\bar{\alpha} > 0$ we see from (39), (41) and (47) that for $t \in (a, b)$

$$u(t)-\bar{\alpha}v(t) > \int_a^b G(s,t)A(s)[u(s)-\bar{\alpha}v(s)]\,ds \geqq 0,$$

hence, (49a) is impossible. Similarly, from (39), (42) and (47)

$$u'(a)-\bar{\alpha}v'(a) > \int_a^b \left(\frac{b-s}{b-a}\right)A(s)[u(s)-\bar{\alpha}v(s)]\,ds \geqq 0,$$

so (49b) is impossible. Finally by virtue of (39), (43) and (47)

$$0 \geqq -\int_a^b \left(\frac{s-a}{b-a}\right)A(s)[u(s)-\alpha v(s)]\,ds > u'(b)-\bar{\alpha}v'(b),$$

which rules out (49c). This contradiction gives the result.

5. Proofs of theorems.

Proof of Theorem 1. Assuming the hypotheses of Theorem 1, let $k > 0$ be so large that every element of the matrix $B(t)+k^2I$ is nonnegative on $[a, b]$. Let σ be the solution of $\sigma'(s) = \exp 2k\sigma(s)$, $\sigma(0) = a$. The function σ will be defined on some interval $[0, d)$ with $\sigma(s) \to +\infty$ as $s \to d$. If $x(t)$ is a nontrivial solution of $x''(t)+B(t)x(t) = 0$ with $x(a) = x(b) = 0$ then $u(s) = x(\sigma(s))e^{-k\sigma(s)}$ satisfies $u''(s)+[B(\sigma(s))+k^2I]e^{4k\sigma(s)}u(s) = 0$, $u(0) = u(\sigma^{-1}(b)) = 0$. Moreover, the elements of $B(\sigma(s))+k^2I$ are nonnegative on $[0, \sigma^{-1}(b)]$. Since $A(t) = (a_{ij}(t))$ is continuous on $[a, b]$ with $a_{ij}(t) \geqq b_{ij}(t)$ and $a_{ij}(\bar{t}) > b_{ij}(\bar{t})$ for $1 \leqq i, j \leqq n$, it follows that every element of $A(\sigma(s))+k^2I$ is greater than or equal to the corresponding element of $B(\sigma(s))+k^2I$ on interval $[0, \sigma^{-1}(b)]$. Also, if $\bar{s} = \sigma^{-1}(\bar{t})$ then every element of $A(\sigma(\bar{s}))+k^2I$ is strictly greater than the corresponding element of $B(\sigma(\bar{s}))+k^2I$. From Lemma 9 we infer the existence of a nontrivial solution $v(s)$ of $v''(s)+[A(\sigma(s))+k^2I]e^{4k\sigma(s)}v(s) = 0$ with $v(0) = v(s^*) = 0$, $0 < s^* < \sigma^{-1}(b)$. Thus, if $y(t) = v(\sigma^{-1}(t))e^{kt}$ then $y(t) \not\equiv 0$, $y''(t)+A(t)y(t) = 0$, $t \in [a, b]$, and $y(a) = y(c) = 0$ where $a = \sigma^{-1}(0) = \sigma^{-1}(s^*) = c < b$. This completes the proof of Theorem 1.

Remark. Making a change of variable to derive an equivalent system, where the diagonal elements of the matrix are also positive, is apparently known. But we have included it in the proof of Theorem 1 for the sake of completeness.

Proof of Theorem 3. By making use of the same device that was used to derive Theorem 1 from Lemma 9 we may assume that if $B(t) = (b_{ij}(t))$ then $b_{ij}(t) \geqq 0$ for $1 \leqq i, j \leqq n$. Assume that the hypotheses of Theorem 3 hold and that there exists a nontrivial solution $y(t)$ of $y''(t)+B(t)y(t) = 0$ which satisfies neither condition (i) nor condition (ii) of the assertion of this theorem. We may assume that $y(t) = \operatorname{col}(y_1(t), \cdots, y_n(t))$ with $y_k(t) > 0$ for $t \in [a, b]$, and $1 \leqq k \leqq n$. Consider

$$v(t) = y(t) - \left[\left(\frac{b-t}{b-a}\right)y(a) + \left(\frac{t-a}{b-a}\right)y(b)\right]. \tag{50}$$

Since $v''(t) = y''(t) = -B(t)y(t)$ and $v(a) = v(b) = 0$ it follows that

$$v(t) = \int_a^b G(s,t)B(s)y(s)\,ds.$$

Therefore, as

$$0 < \left[\left(\frac{b-t}{b-a}\right)y(a) + \left(\frac{t-a}{b-a}\right)y(b)\right],$$

we have the inequality

$$\int_a^b G(s, t)B(s)y(s)\,ds < y(t) \tag{51}$$

for $t \in [a, b]$.

If $x(t)$ is a nontrivial solution of the boundary value problem $x''(t)+B(t)x(t)=0$ with $x(a)=x(b)=0$ then

$$x(t)=\int_a^b G(s, t)B(s)x(s)\,ds. \tag{52}$$

If $x(t)=\text{col}\,(x_1(t), \cdots, x_n(t))$ we may assume that there exist l and $\bar{t} \in [a, b]$ with $x_l(\bar{t})>0$. It follows that if $\alpha>0$ and sufficiently small then

$$0<y(t)-\alpha x(t), \qquad t \in [a, b], \tag{53}$$

while if α is sufficiently large, some component of $y(t)-\alpha x(t)$ must vanish somewhere on $[a, b]$. Thus if $\bar{\alpha}$ is the least upper bound of the set of numbers α such that (53) holds everywhere on $[a, b]$ then

$$0 \leqq y(t)-\bar{\alpha}x(t), \tag{54}$$

and

$$0=y_k(c)-\bar{\alpha}x_k(c) \tag{55}$$

for some k, and some $c \in (a, b)$. From (51), (52), (54), and the assumption that the elements of $B(t)$ are nonnegative we see that

$$0=y_k(c)-\bar{\alpha}x_k(c)>\int_a^b G(s, t) \sum_{j=1}^n b_{kj}(s)[y_j(s)-\bar{\alpha}x_j(s)]\,ds \geqq 0.$$

This contradiction proves the theorem.

Proof of Theorem 4. By using the same device that allowed us to derive Theorem 1 from Lemma 9 we may again assume that if $B(t)=(b_{ij}(t))$ then $b_{ij}(t) \geqq 0$ for $1 \leqq i, j \leqq n$ and $t \in [a, b]$. For each integer $m=1, 2, \cdots$, let $B_m(t)=(b_{ij}(t)+1/m)$. Let $x(t)$ be a nontrivial solution of the boundary value problem $x''(t)+B(t)x(t)=0$, $x(a)=x(b)=0$, and assume there exists no nontrivial solution of the boundary value problem $x''(t)+B(t)x(t)=0$, $x(a)=x(c)=0$ if $a<c<b$. As every element of $B_m(t)$ is strictly greater than the corresponding element of $B(t)$, it follows from Lemma 9 that there exists a nontrivial solution of the boundary value problem $u_m''(t)+B_m(t)u_m(t)=0$, $u_m(a)=u_m(c_m)=0$, such that $a<c_m<b$ and such that $u_m(t) \in K(c_m)$. As

$$u_m(t)=\int_a^{c_m} G(t, s, c_m)B_m(s)u_m(s)\,ds,$$

for $a \leqq t \leqq c_m$, the same argument that was used to establish the inequality (12) shows that

$$1 \geqq \frac{4}{(c_m-a)\int_a^{c_m} \|B_m(s)\|\,ds}.$$

Thus, since $\|B_m(t)\|=n/m+\|B(s)\|$ is bounded independently of m, we infer the existence of a number $\delta>0$ such that

$$a+\delta \leqq c_m<b, \qquad m \geqq 1. \tag{56}$$

As in the proof of Lemma 9 we may assume, without loss of generality, that $u'_m(a)\to k\neq 0$ as $m\to\infty$ and that $\lim_{m\to\infty} c_m = c$ with $a+\delta \leqq c \leqq b$. If $u''(t)+B(t)u(t)=0$, $u(a)=0$ and $u'(a)=k$ then the sequence $\{u_m(t)\}_1^\infty$ converges uniformly to $u(t)$ on $[a, b]$ and hence $u(c)=0$. If $c<b$ we would have a contradiction to the previous assumption concerning b. If $a<\bar{t}<b$ then $\bar{t}<c_m$ for sufficiently large m and, as $u_m \in K(c_m)$, $0\leqq u_m(\bar{t})$. Hence $0\leqq u(\bar{t})$, so $u\in K(b)$ and the theorem is proved.

Proof of Theorem 2. The proof of Theorem 2 will follow quickly from Theorems 1 and 3 and part of the above argument. As before, we may suppose that elements of $B(t)$ are nonnegative on $[a, b]$. For each $m\geqq 1$, let $B_m(t)$ have the same meaning as in the proof of Theorem 3. Assume that $x(t)$ is a nontrivial solution of $x''(t)+B(t)x(t)=0$, $x(a)=x(b)=0$ and there is no nontrivial solution of the differential equation vanishing at a and c if $a<c<b$. Let $a<t_1<t_2<b$. Referring to the proof of Theorem 4 we see that there is an r such that there exists $u_r(t)\not\equiv 0$ satisfying $u_r''(t)+B_r(t)u_r(t)=0$, $u_r(a)=u_r(c_r)=0$, with $t_2<c_r<b$, and such that $0\leqq u_r(t)$ for $a<t<c_r$. We assert that

$$0<u_r(t), \qquad t\in[t_1, t_2]. \tag{57}$$

Indeed, this follows from the equation

$$u_r(t)=\int_a^{c_r} G(s, t, c_r)B_r(s)u_r(s)\,ds,$$

the strict positivity of the elements of B_r and (8). Suppose contrary to the assertion of Theorem 2 there exists $y(t)\not\equiv 0$ such that $y''(t)+B(t)y(t)=0$ and $y(t_1)=y(t_2)=0$. As every element of $B_r(t)$ is greater than the corresponding element of $B(t)$ it follows from Theorem 1 that there exists $z(t)\not\equiv 0$ such that $z''(t)+B_r(t)z(t)=0$ and $z(t_1)=z(d)=0$ where $t_1<d<t_2$. Since $u_r(t)$ satisfies the same differential equation as $z(t)$ and since every element of $u_r(t)$ is strictly positive on $[t_1, d]$, by (57), this contradicts the statement of Theorem 3. This contradiction proves Theorem 2.

Remark. The results of this paper were announced in the Notices Amer. Math. Soc., 23 (1976), no. 7, 76T-B205.

REFERENCES

[1] S. Ahmad and A. C. Lazer, *On the components of extremal solutions of second order systems*, this Journal, 8 (1977), pp. 16–23.

[2] R. Bellman, *Matrix Analysis*, 2nd ed., McGraw-Hill, New York, 1970.

[3] G. Birkhoff and M. Hestenes, *Natural isoperimetric conditions in the calculus of variations*, Duke Math. J., 1 (1935), pp. 198–286.

[4] W. Coppel, *Disconjugacy*, Lecture Notes in Mathematics 20, Springer-Verlag, Berlin–Heidelberg–New York, 1971.

[5] M. Krein and M. Rutman, *Linear operators leaving invariant a cone in a Banach space*, Uspehi Math. Nauk. 3 (1948), no. 1 (23), pp. 3–95. (In Russian; English translation: American Math. Soc. Transl., no. 26.)

[6] M. Morse, *A generalization of the Sturm separation and comparison theorems in n-space*, Math. Ann. 103 (1930), pp. 52–69.

[7] ———, *Variational Analysis: Critical Extremals and Sturmian Extensions*, John Wiley, New York, 1973.

[8] W. T. Reid, *A comparison theorem for selfadjoint differential equations of second order*, Ann. Math. 65 (1957), pp. 197–202.

[9] ———, *Ordinary Differential Equations*, John Wiley, New York, 1971.

[10] J. C. F. Sturm, *Mémoire sur les équations différentielles linéaires de second ordre*, J. Math. Pures Appl. 1 (1836), pp. 106–186.

SIAM J. MATH. ANAL.
Vol. 9, No. 6, December 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0014 $01.00/0

NUMERICAL APPROXIMATION OF NONLINEAR FUNCTIONAL DIFFERENTIAL EQUATIONS WITH L^2 INITIAL FUNCTIONS*

G. W. REDDIEN† AND G. F. WEBB†

Abstract. Nonlinear operator semigroup theory is used to treat the numerical approximation of autonomous functional differential equations with L^2 initial functions. The consistency, stability, and convergence of both explicit and implicit schemes are demonstrated and error estimates are established. The stability is obtained easily as a consequence of a renorming of the underlying space.

1. Introduction. The objective of this paper is to study the numerical approximation of solutions to nonlinear autonomous functional differential equations of the form

$$\text{(FDE)}\qquad \begin{aligned} &x_0=\phi\in L^2(-r,0;R^n), \quad x(0)=h\in R^n,\\ &\dot{x}(t)=F(x_t), \qquad t\geqq 0.\end{aligned}$$

The notation of (FDE) means r is a fixed positive number $F: L^2(-r,0;R^n)\to R^n$, and for $t\geqq 0$, $x_t\in L^2(-r,0;R^n)$ is defined by $x_t(\theta)=x(t+\theta)$ for almost all $\theta\in[-r,0]$. It is natural to choose $L^2(-r,0;R^n)$ as a space of initial functions for (FDE) and several authors who have studied (FDE) in this space are listed in our references. Not only does $L^2(-r,0;R^n)$ give a more general class of initial functions than $C(-r,0;R^n)$, but its topological and algebraic structure is very advantageous for certain problems. For example, the weak pre-compactness of bounded sets in $L^2(-r,0;R^n)$ is useful in control theory and the simplicity of adjoint operators in $L^2(-r,0;R^n)$ is useful in stability theory. The numerical study of the linear version of (FDE) was first undertaken by H. T. Banks and J. A. Burns [1] and H. T. Banks and A. Manitius [2]. Our treatment of the nonlinear version will be similar to the linear treatment in [1] and will be based upon the existence theory for these equations developed in G. F. Webb [15].

Our method will use a nonlinear operator approach and our problem will be studied in the setting of semigroups of nonlinear operators. The advantage of this approach is to greatly simplify the analytical details in demonstrating the convergence of the approximating solutions, as well as clarify error estimates in the rate of convergence. In demonstrating the convergence of approximations to most types of evolution equations there are two fundamental problems to be resolved. One is the consistency of the approximating scheme, which is usually very easy. The other is the stability of the approximating scheme, which is usually very difficult. The main contribution of this paper is to formulate our evolution equation in a specially chosen space so that both the consistency and stability of our approximating schemes are easily and naturally obtained.

We set forth below some needed facts from the general theory of semigroups of nonlinear operators in Hilbert space (see M. G. Crandall and A. Pazy [7]).

(1.1) A semigroup of nonlinear operators $T(t)$, $t\geqq 0$ in a Hilbert space X is a family of Lipschitz continuous operators in X satisfying $T(0)=I$, $T(t+s)=T(t)T(s)$ for $s,t\geqq 0$, $T(t)x$ is continuous from $[0,\infty)$ to X for each fixed $x\in X$, and there exists a real constant γ such that $\|T(t)x-T(t)y\|\leqq e^{\gamma t}\|x-y\|$ for all $x,y\in X$ and $t\geqq 0$.

* Received by the editors May 24, 1976, and in final revised form July 1, 1977.
† Department of Mathematics, Vanderbilt University, Nashville, Tennessee 37235.

(1.2) The operator $A: X \to X$ is the generator of $T(t), t \geqq 0$ provided $\lim_{n \to \infty} (I+(t/n)A)^{-n}x = T(t)x$ for all $x \in X$ and $t \geqq 0$.

(1.3) If $x \in D(A)$, then $T(t)x \in D(A)$ and $\|AT(t)x\| \leqq e^{\gamma t}\|Ax\|$ for all $t \geqq 0$.

(1.4) If $x \in D(A)$, then $\|T(t)x - T(s)x\| \leqq |t-s|\, e^{|\gamma| t}\|Ax\|$ for all $0 \leqq s \leqq t$.

(1.5) If $x \in D(A)$, then $d^+/dt\, T(x) = -AT(t)x$ for all $t \geqq 0$.

In addition to the facts stated above we will assume throughout this paper the results and notation of [15].

As in [15] we will consider two separate classes of equations, which we describe as the continuous case (which is easy) and the discontinuous case (which is more difficult, but which includes the important equations of delay type). The continuous case will be discussed in § 2 and the discontinuous case in § 3. For each case, we will consider an explicit and implicit approximation scheme. In § 4 we will give some numerical examples to illustrate the computational efficacy of our development.

2. The continuous case. In this case we treat (FDE) for F Lipschitz continuous from $L^2(-r, 0; R^n)$ to R^n. We let β be the Lipschitz constant for F. From [15] we have associated with the solutions of (FDE) a semigroup of nonlinear operators $T(t)$, $t \geqq 0$ acting in the space $X = L^2(-r, 0; R^n) \times R^n$ with inner product and norm

$$\langle \{\phi, h\}, \{\psi, k\} \rangle = \int_{-r}^{0} (\phi(\theta), \psi(\theta))\, d\theta + (h, k), \tag{2.1}$$
$$\|\{\phi, h\}\|^2 = \int_{-r}^{0} |\phi(\theta)|^2\, d\theta + |h|^2.$$

As in [15] we define two projections π_1 and π_2 in X by $\pi_1\{\phi, h\} = \phi$ and $\pi_2\{\phi, h\} = h$. For any $\{\phi, h\}$ in X the solution of (FDE) is identified with the semigroup $T(t)$, $t \geqq 0$, by means of the formulas

$$x(\phi, h)(t) = \pi_2 T(t)\{\phi, h\}, \qquad t \geqq 0, \tag{2.2}$$
$$x_t(\phi, h) = \pi_1 T(t)\{\phi, h\}, \qquad t \geqq 0,$$

[15, prop. 5.9]. The generator of $T(t)$, $t \geqq 0$, is $A: X \to X$

$$D(A) = \{\{\phi, h\} : \phi \text{ is absolutely continuous, } \phi' \in L^2(-r, 0; R^n),\ \phi(0) = h\}, \tag{2.3}$$
$$A\{\phi, h\} = \{-\phi', -F(\phi)\}$$

[15, Props. 3.1, 3.2]. The constant γ in (1.1), (1.3), and (1.4) is given by $\gamma = \beta + \frac{1}{2}$ [15, Prop. 3.1].

Our approximations to the solutions of (FDE) will be defined using the following "averaging" projections: Let N be a positive integer, let $\chi_i^N = \chi_{[-ri/N, -r(i-1)/N)}$ for $i = 1, \cdots, N$, and let X_N be the subspace of X defined by

$$X_N = \left\{ \{\phi, h\} \in X : \phi = \sum_{i=1}^{N} h_i \chi_i^N,\ h_i \in R^n \right\}.$$

As in [1] we define $P_N: X \to X_N$ by

$$P_N\{\phi, h\} = \left\{ \sum_{i=1}^{N} h_i \chi_i^N, h \right\}, \quad \text{where } h_i = (N/r) \int_{-ri/N}^{-r(i-1)/N} \phi(\theta)\, d\theta. \tag{2.4}$$

We observe that for $\{\phi, h\} \in X_N$, $\|\{\phi, h\}\|^2 = (r/N) \sum_{i=1}^N |h_i|^2 + |h|^2$. Also, for $\{\phi, h\} \in X$, $\|P_N\{\phi, h\}\| \leqq \|\{\phi, h\}\|$, so that the P_N's are uniformly bounded in the operator norm. For each N let $F_N: \pi_1 X_N \to R^n$ Lipschitz continuously with Lipschitz constant β_N and let

$$\sup_{N \geqq 1} \beta_N < \infty \tag{2.5}$$

$$\lim_{N\to\infty} F_N(\pi_1 P_N\{\phi, h\}) = F(\phi) \quad \text{for all } \{\phi, h\} \in X. \tag{2.6}$$

We note that (2.6) holds uniformly on compact subsets of X by virtue of (2.5). We define a sequence of discretized approximations to A by

$$\begin{aligned} D(A_N) = X_N \quad \text{and} \quad &\text{for } \{\phi, h\} = \left\{ \sum_{i=1}^N h_i \chi_i^N, h_0 \right\}, \\ &A_N\{\phi, h\} = \left\{ -\sum_{i=1}^N (N/r)(h_{i-1} - h_i)\chi_i^N, -F_N(\phi) \right\}. \end{aligned} \tag{2.7}$$

Our explicit approximation result is given by

THEOREM 2.1. *If $\{t_N\}_{N=1}^\infty$ is a decreasing sequence of positive numbers such that $t_N \leqq r/N$, $\{\phi, h\} \in X$, and $t \geqq 0$, then*

$$\lim_{N\to\infty} \|(I - t_N A_N)^{[t/t_N]} P_N\{\phi, h\} - P_N T(t)\{\phi, h\}\| = 0 \tag{2.8}$$

(here $[t/t_N]$ denotes the greatest integer $\leqq t/t_N$).

Our implicit approximation result is given by

THEOREM 2.2. *If $\{t_N\}_{N=1}^\infty$ is a decreasing sequence of positive numbers converging to 0, such that $t_N < (\beta_N + 1/2)^{-1}$, $\{\phi, h\} \in X$,*

$$\lim_{N\to\infty} \|(I + t_N A_N)^{-[t/t_N]} P_N\{\phi, h\} - P_N T(t)\{\phi, h\}\| = 0. \tag{2.9}$$

Before giving the proof of Theorems 2.1 and 2.2 we illustrate how one defines the F_N's for some examples.

Let Lip (R^n, R^n) denote the Banach space of Lipschitz continuous functions from R^n to R^n with norm $\|f\|_{\text{Lip}} = |f|_{\text{Lip}} + |f(0)|$, where $|f|_{\text{Lip}}$ is the Lipschitz constant for f, that is

$$|f|_{\text{Lip}} = \sup_{x,y \in R^n, x \neq y} |f(x) - f(y)|/|x - y|.$$

Example 2.1. Let $f, g \in \text{Lip}(R^n, R^n)$ and consider the equation

$$\dot{x}(t) = g\left(\int_{t-r}^t f(x(s))\, ds \right), \quad x_0 = \phi \in L^2(-r, 0; R^n), \quad x(0) = h \in R^n.$$

Let $F(\phi) \overset{\text{def}}{=} g(\int_{-r}^0 f(\phi(s))\, ds)$ and let $\{f_N\}_{N=1}^\infty$ and $\{g_N\}_{N=1}^\infty$ be sequences in Lip(R^n, R^n) converging to f and g, respectively (in § 4 we give some specific examples for f, g, f_N, g_N). Define $F_N(\sum_{i=1}^N h_i \chi_i^N) = g_N((r/N) \sum_{i=1}^N f_N(h_i))$. To see that (2.5) is satisfied observe that

$$\left|F_N\Big(\sum_{i=1}^N h_i\chi_i^N\Big)-F_N\Big(\sum_{i=1}^N k_i\chi_i^N\Big)\right|$$

$$\leqq |g_N|_{\text{Lip}}|f_N|_{\text{Lip}}\frac{r}{N}\sum_{i=1}^N |h_i-k_i|$$

$$\leqq (\sup_{N\geqq 1}|g_N|_{\text{Lip}})(\sup_{N\geqq 1}|f_N|_{\text{Lip}})r^{1/2}\Big(\frac{r}{N}\sum_{i=1}^N |h_i-k_i|^2\Big)^{1/2}.$$

To see that (2.6) is satisfied observe that for $\{\phi, h\}\in X$ such that ϕ is continuous

$$|F_N(\pi_1 P_N\{\phi, h\})-F(\phi)|$$

$$=\left|g_N\Big(\frac{r}{N}\sum_{i=1}^N f_N\Big(\frac{N}{r}\int_{-ri/N}^{-r(i-1)/N}\phi(s)\,ds\Big)\Big)-g\Big(\sum_{i=1}^N\int_{-ri/N}^{-r(i-1)/N} f(\phi(s))\,ds\Big)\right|$$

$$\leqq\left|g_N\Big(\frac{r}{N}\sum_{i=1}^N f_N\Big(\frac{N}{r}\int_{-ri/N}^{-r(i-1)/N}\phi(s)\,ds\Big)\Big)-g_N\Big(\frac{r}{N}\sum_{i=1}^N f\Big(\frac{N}{r}\int_{-ri/N}^{-r(i-1)/N}\phi(s)\,ds\Big)\Big)\right|$$

$$+\left|g_N\Big(\frac{r}{N}\sum_{i=1}^N f\Big(\frac{N}{r}\int_{-ri/N}^{-r(i-1)/N}\phi(s)\,ds\Big)\Big)-g\Big(\frac{r}{N}\sum_{i=1}^N f\Big(\frac{N}{r}\int_{-ri/N}^{-r(i-1)/N}\phi(s)\,ds\Big)\Big)\right|$$

$$+\left|g\Big(\frac{r}{N}\sum_{i=1}^N f\Big(\frac{N}{r}\int_{-ri/N}^{-r(i-1)/N}\phi(s)\,ds\Big)\Big)-g\Big(\sum_{i=1}^N\int_{-ri/N}^{-r(i-1)/N} f(\phi(s))\,ds\Big)\right|$$

$$\leqq |g_N|_{\text{Lip}}\frac{r}{N}\sum_{i=1}^N\Big(|f_N-f|_{\text{Lip}}\Big|\frac{N}{r}\int_{-ri/N}^{-r(i-1)/N}\phi(s)\,ds\Big|+|(f_N-f)(0)|\Big)$$

$$+|g_N-g|_{\text{Lip}}\frac{r}{N}\sum_{i=1}^N|f(\phi(s_i))|+|(g_N-g)(0)|$$

$$+|g|_{\text{Lip}}\sum_{i=1}^N\Big|\frac{r}{N}f(\phi(s_i))-\int_{-ri/N}^{-r(i-1)/N}f(\phi(s))\,ds\Big|$$

$$\Big(\text{where } s_i\in[-ri/N, -r(i-1)/N] \text{ such that } \phi(s_i)=\frac{N}{r}\int_{-ri/N}^{-r(i-1)/N}\phi(s)\,ds\Big)$$

$$\leqq |g_N|_{\text{Lip}}\Big(|f_N-f|_{\text{Lip}}\int_{-r}^{0}|\phi(s)|\,ds+r|(f_N-f)(0)|\Big)$$

$$+|g_N-g|_{\text{Lip}}r\sup_{1\leqq i\leqq N}|f(\phi(s_i))|+|(g_N-g)(0)|$$

$$+|g|_{\text{Lip}}|f|_{\text{Lip}}\sum_{i=1}^N\int_{-ri/N}^{-r(i-1)/N}|\phi(s_i)-\phi(s)|\,ds.$$

Thus, (2.6) holds for $\{\phi, h\}$ such that ϕ is continuous. Since these $\{\phi, h\}$ are dense in X, F is Lipschitz continuous, and the F_N's are uniformly Lipschitz continuous in N, (2.6) holds for all $\{\phi, h\}\in X$.

Before proving Theorems 2.1 and 2.2 we establish some lemmas. The first lemma gives the stability of the explicit approximating scheme.

LEMMA 2.1. *For each N let* $\gamma_N = \beta_N + \frac{1}{2}$. *If* $0 < \lambda \leqq r/N$ *and* $\{\phi, h\}, \{\psi, k\} \in X_N$, *then*

$$\|(I - \lambda A_N)\{\phi, h\} - (I - \lambda A_N)\{\psi, k\}\| \leqq (1 + \lambda \gamma_N)\|\{\phi, h\} - \{\psi, k\}\|. \tag{2.10}$$

Proof. For $\{\phi, h\}, \{\psi, k\} \in X_N$, $\phi = \sum_{i=1}^N h_i \chi_i^N$, $h = h_0$, $\psi = \sum_{i=1}^N k_i \chi_i^N$, $k = k_0$, we have

$$\begin{aligned}
&\|(I - \lambda A_N)\{\phi, h\} - (I - \lambda A_N)\{\psi, k\}\| \\
&= \left\| \left\{ \sum_{i=1}^N \left(\left(1 - \frac{\lambda N}{r}\right)(h_i - k_i) + \left(\frac{\lambda N}{r}\right)(h_{i-1} - k_{i-1}) \right) \chi_i^N, h - k + \lambda (F_N(\phi) - F_N(\psi)) \right\} \right\| \\
&\leqq \left(1 - \frac{\lambda N}{r}\right) \left\| \left\{ \sum_{i=1}^N (h_i - k_i) \chi_i^N, h - k \right\} \right\| \\
&\quad + \left(\frac{\lambda N}{r}\right) \left\| \left\{ \sum_{i=1}^N (h_{i-1} - k_{i-1}) \chi_i^N, h - k + \frac{r}{N}(F_N(\phi) - F_N(\psi)) \right\} \right\|.
\end{aligned}$$

Then, (2.10) follows with the use of the inequality

$$\begin{aligned}
&\left\| \left\{ \sum_{i=1}^N (h_{i-1} - k_{i-1}) \chi_i^N, h_0 + \frac{r}{N} F_N(\phi) - k_0 - \frac{r}{N} F_N(\psi) \right\} \right\|^2 \\
&= \frac{r}{N} \sum_{i=1}^N |h_{i-1} - k_{i-1}|^2 + |h_0 - k_0|^2 \\
&\quad + (2r/N)(h_0 - k_0, F_N(\phi) - F_N(\psi)) + (r/N)^2 |F_N(\phi) - F_N(\psi)|^2 \\
&\leqq \|\{\phi, h\} - \{\psi, k\}\|^2 + (r/N)|h_0 - k_0|^2 \\
&\quad + (r/N)\beta_N(|h_0 - k_0|^2 + |\phi - \psi|^2) + (r/N)^2 \beta_N^2 |\phi - \psi|^2 \\
&\leqq (1 + (r/N)\gamma_N)^2 \|\{\phi, h\} - \{\psi, k\}\|^2.
\end{aligned}$$

The next lemma establishes the stability of the implicit approximating scheme.

LEMMA 2.2. *For each N let* $\gamma_N = \beta_N + \frac{1}{2}$. *Then,* $A_N + \gamma_N I$ *is accretive in* X_N. *Consequently, for* $0 < \lambda < 1/\gamma_N$, $(I + \lambda A_N)^{-1}$ *exists, is everywhere defined on* X_N, *and satisfies for* $\{\phi, h\}, \{\psi, k\} \in X_N$.

$$\|(I + \lambda A_N)^{-1}\{\phi, h\} - (I + \lambda A_N)^{-1}\{\psi, k\}\| \leqq (1 - \lambda \gamma_N)^{-1} \|\{\phi, h\} - \{\psi, k\}\|. \tag{2.11}$$

Proof. Let $\{\phi, h\}, \{\psi, k\} \in X_N$, $\phi = \sum_{i=1}^N h_i \chi_i^N$, $h = h_0$, $\psi = \sum_{i=1}^N k_i \chi_i^N$, $k = k_0$, and $u_i = h_i - k_i$, $i = 0, 1, \cdots, N$. The accretiveness of $A_N + \gamma_N I$ is equivalent to the following inequality:

$$\begin{aligned}
&\langle A_N\{\phi, h\} - A_N\{\psi, k\}, \{\phi, h\} - \{\psi, k\} \rangle \\
&= \left\langle \left\{ \sum_{i=1}^N -\frac{N}{r}(u_{i-1} - u_i)\chi_i^N, -F_N(\phi) + F_N(\psi) \right\}, \left\{ \sum_{i=1}^N u_i \chi_i^N, u_0 \right\} \right\rangle \\
&= -\sum_{i=1}^N (u_{i-1} - u_i, u_i) - (F_N(\phi) - F_N(\psi), u_0) \\
&= \sum_{i=1}^N \tfrac{1}{2}|u_{i-1} - u_i|^2 - \tfrac{1}{2}|u_0|^2 + \tfrac{1}{2}|u_N|^2 - (F_N(\phi) - F_N(\psi), u_0) \\
&\geqq -\tfrac{1}{2}|u_0|^2 - \beta_N |\phi - \psi|\, |u_0| \\
&\geqq -\gamma_N \|\{\phi, h\} - \{\psi, k\}\|^2.
\end{aligned}$$

The last statement of the lemma follows from the accretiveness of $A_N+\gamma_N I$ and the continuity of A_N (see [14]).

The next lemma together with Lemma 2.4 will establish the consistency of both explicit and implicit approximating schemes.

LEMMA 2.3. *If* $\{\phi, h\}\in D(A)$ *and* $t\geqq 0$, *then*

$$\begin{aligned}\|P_N AT(t)\{\phi, h\}-A_N P_N T(t)\{\phi, h\}\|^2 &\\ \leqq \int_{-r/N}^{0} |(\pi_1 AT(t)\{\phi, h\})(\theta)|^2 d\theta &\\ +\frac{1}{3}\int_{-r}^{-r/N} \left|(\pi_1 AT(t)\{\phi, h\})(\theta)-(\pi_1 AT(t)\{\phi, h\})\left(\theta+\frac{r}{N}\right)\right|^2 d\theta &\\ +|F(\pi_1 T(t)\{\phi, h\})-F_N(\pi_1 P_N T(t)\{\phi, h\})|^2.&\end{aligned}\tag{2.12}$$

Proof. Let $\{\psi, k\}=T(t)\{\phi, h\}$ and for $i=1,\cdots,N$ let

$$v_i=\frac{N}{r}\int_{-ri/N}^{-r(i-1)/N}\psi(s)\,ds,\qquad v_0=k,$$

$$w_i=\frac{N}{r}\int_{-ri/N}^{-r(i-1)/N}\psi'(s)\,ds=(N/r)(\psi(-r(i-1)/N)-\psi(-ri/N)).$$

By (1.3), $\{\psi, k\}\in D(A)$ and we have

$$\begin{aligned}\|P_N A\{\psi, k\}-A_N P_N\{\psi, k\}\|^2 &\\ =\|P_N\{-\psi', -F(\psi)\}-A_N P_N\{\psi, k\}\|^2 &\\ =\left\|\left\{\sum_{i=1}^N w_i\chi_i^N, F(\psi)\right\}-\left\{\sum_{i=1}^N \frac{N}{r}(v_{i-1}-v_i)\chi_i^N, F_N(\pi_1 P_N\{\psi, k\})\right\}\right\|^2 &\\ =\frac{N}{r}\sum_{i=1}^N\left|\psi\left(\frac{-r(i-1)}{N}\right)-\psi\left(\frac{-ri}{N}\right)-(v_{i-1}-v_i)\right|^2+|F(\psi)-F_N(\pi_1 P_N\{\psi, k\})|^2.&\end{aligned}\tag{2.13}$$

Since $\{\psi, k\}\in D(A)$, $\psi(0)=k=v_0$, so that for $i=1$ we have

$$\begin{aligned}|\psi(0)-\psi(-r/N)-(v_0-v_1)| &=\left|\psi\left(\frac{-r}{N}\right)-\frac{N}{r}\int_{-r/N}^0\psi(s)\,ds\right|\\ &\leqq\frac{N}{r}\int_{-r/N}^0|\psi(-r/N)-\psi(s)|ds\leqq\frac{N}{r}\int_{-r/N}^0\int_{-r/N}^s|\psi'(u)|\,du\,ds\\ &\leqq\int_{-r/N}^0|\psi'(u)|\,du\\ &\leqq\left(\frac{r}{N}\right)^{1/2}\left(\int_{-r/N}^0|\psi'(u)|^2\,du\right)^{1/2}\\ &=\left(\frac{r}{N}\right)^{1/2}\left(\int_{-r/N}^0|(\pi_1 AT(t)\{\phi, h\})(\theta)|^2\,d\theta\right)^{1/2}.\end{aligned}$$

For $i=2,\cdots,N$

$$
\begin{aligned}
&|\psi(-r(i-1)/N)-\psi(-ri/N)-(v_{i-1}-v_i)|\\
&=\left|\psi\left(\frac{-r(i-1)}{N}\right)-\psi\left(\frac{-ri}{N}\right)-\frac{N}{r}\int_{-ri/N}^{-r(i-1)/N}\left(\psi\left(s+\frac{r}{N}\right)-\psi(s)\right)ds\right|\\
&=\left|\psi\left(\frac{-r(i-1)}{N}\right)-\psi\left(\frac{-ri}{N}\right)-\frac{N}{r}\left(-\int_{-ri/N}^{-r(i-1)/N}\left(\left(\frac{r(i-1)}{N}\right)+s\right)\left(\psi'\left(s+\frac{r}{N}\right)-\psi'(s)\right)ds\right.\right.\\
&\quad\left.\left.+(0)\left(\psi\left(\frac{-r(i-2)}{N}\right)-\psi\left(\frac{-r(i-1)}{N}\right)\right)-\left(\frac{-r}{N}\right)\left(\psi\left(\frac{-r(i-1)}{N}\right)-\psi\left(\frac{-ri}{N}\right)\right)\right)\right|\\
&=\left|\frac{N}{r}\int_{-ri/N}^{-r(i-1)/N}\left(\left(\frac{r(i-1)}{N}\right)+s\right)\left(\psi'\left(s+\frac{r}{N}\right)-\psi'(s)\right)ds\right|\\
&\leqq\frac{N}{r}\left(\int_{-ri/N}^{-r(i-1)/N}\left(\left(\frac{r(i-1)}{N}\right)+s\right)^2ds\right)^{1/2}\left(\int_{-ri/N}^{-r(i-1)/N}\left|\psi'\left(s+\frac{r}{N}\right)-\psi'(s)\right|^2ds\right)^{1/2}\\
&=\left(\frac{r}{3N}\right)^{1/2}\left(\int_{-ri/N}^{-r(i-1)/N}\left|\psi'(s)-\psi'\left(s+\frac{r}{N}\right)\right|^2ds\right)^{1/2}\\
&=\left(\frac{r}{3N}\right)^{1/2}\left(\int_{-ri/N}^{-r(i-1)/N}\left|(\pi_1AT(t)\{\phi,h\})(\theta)-(\pi_1AT(t)\{\phi,h\})\left(\theta+\frac{r}{N}\right)\right|^2d\theta\right)^{1/2}.
\end{aligned}
$$

The conclusion (2.12) now follows.

LEMMA 2.4. *If* $\{\phi, h\}\in D(A)$, *then*

(2.14) $t\to AT(t)\{\phi,h\}$ *is continuous from* $[0,\infty)$ *to* X;

(2.15) *for* $t\geqq 0$, $|u|<\delta$,

$$\|T(t+u)\{\phi,h\}-T(t)\{\phi,h\}+uAT(t)\{\phi,h\}\|$$

$$\leqq\delta\max_{t-\delta\leqq s\leqq t+\delta}\|AT(s)\{\phi,h\}-AT(t)\{\phi,h\}\|;$$

(2.16) $\dot{x}(\phi,h)(t)=\phi'(t)$ *for a.e.* $t\in[-r,0]$ *and* $F(x_t(\phi,h))$ *for* $t\geqq 0$;

(2.17) $AT(t)\{\phi,h\}=\{(x_t(\phi,h))', \dot{x}(\phi,h)(t)\}$ *for* $t\geqq 0$.

Proof. (2.16) and (2.17) follow from (1.3), (2.2), (2.3), and Proposition 5.9 of [15]. To prove (2.14) we observe that F Lipschitz continuous from $L^2(-r,0;R^n)$ to R^n and $t\to x_t(\phi,h)=\pi_1T(t)\{\phi,h\}$ continuous from $[0,\infty)$ to $L^2(-r,0;R^n)$ implies that $t\to\dot{x}(\phi,h)(t)$ is continuous from $[0,\infty)$ to R^n. The continuity of the mapping $t\to(x_t(\phi,h))'$ from $[0,\infty)$ to $L^2(-r,0;R^n)$ then follows from the fact that $(x_t(\phi,h))'(\theta)=\dot{x}(\phi,h)(t+\theta)$, $\phi'\in L^2(-r,0,R^n)$, and the continuity of translation in $L^2(-r,0;R^n)$. To prove (2.15) one uses (1.5), (2.14), the fact that a continuous right derivative implies the existence of a continuous two sided derivative, and the fundamental theorem of calculus.

Proof of Theorem 2.1. Since $D(A)$ is dense in X, $|T(t)|_{\mathrm{Lip}}\leqq e^{\gamma t}$, and $|(I-t_NA_N)^{[t/t_N]}|_{\mathrm{Lip}}\leqq e^{t\gamma_N}$, it suffices to show (2.8) for $\{\phi,h\}\in D(A)$. Let $\{\phi,h\}\in D(A)$, $t\geqq 0$, and $m=[t/t_N]$. From (1.4) we have

(2.18) $$\|P_NT(t)\{\phi,h\}-P_NT(mt_N)\{\phi,h\}\|\leqq t_Ne^{\gamma t}\|A\{\phi,h\}\|.$$

From (2.12) and (2.17) we have for $k=1, \cdots, m$

$$\begin{aligned}
&\|P_NAT((k-1)t_N)\{\phi, h\}-A_NP_NT((k-1)t_N)\{\phi, h\}\|^2 \\
&\quad \leqq \int_{-r/N}^{0} |\dot{x}(\phi, h)((k-1)t_N+\theta)|^2\,d\theta \\
&\qquad +\frac{1}{3}\int_{-r}^{0}\left|\dot{x}(\phi, h)((k-1)t_N+\theta)-x(\phi, h)\left((k-1)t_N+\theta+\frac{r}{N}\right)\right|^2 d\theta \\
&\qquad +|F(\pi_1 T(k-1)t_N)\{\phi, h\})-F_N(\pi_1 P_N T((k-1)t_N)\{\phi, h\})|^2.
\end{aligned} \tag{2.19}$$

The right side of (2.19) can be made arbitrarily small for all sufficiently large N by Lemma 2.4 and (2.6) (where we use the fact that (2.6) holds uniformly on compact sets in X). Then (2.8) follows from Lemmas 2.1, 2.4, and

$$\begin{aligned}
&\|P_NT(mt_N)\{\phi, h\}-(I-t_NA_N)^mP_N\{\phi, h\}\| \\
&\quad =\left\|\sum_{k=1}^{m}(I-t_NA_N)^{m-k}P_NT(t_N)^k\{\phi, h\}\right. \\
&\qquad \left. -(I-t_NA_N)^{m-k+1}P_NT(t_N)^{k-1}\{\phi, h\}\right\| \\
&\quad \leqq \sum_{k=1}^{m}(1+t_N\gamma_N)^{m-k}\|(I-t_NA_N)P_NT((k-1)t_N)\{\phi, h\}-P_NT(kt_N)\{\phi, h\}\| \\
&\quad \leqq e^{t\gamma_N}\sum_{k=1}^{m}(\|P_NT(kt_N)\{\phi, h\}-P_NT((k-1)t_N)\{\phi, h\} \\
&\qquad\qquad +t_NP_NAT((k-1)t_N)\{\phi, h\}\| \\
&\qquad\qquad +t_N\|P_NAT((k-1)t_N)\{\phi, h\}-A_NP_NT((k-1)t_N)\{\phi, h\}\|).
\end{aligned} \tag{2.20}$$

Proof of Theorem 2.2. The proof is the same as the proof of Theorem 2.1 except that we use Lemma 2.2 to obtain

$$\begin{aligned}
&\|P_NT(mt_N)\{\phi, h\}-(I+t_NA_N)^{-m}P_N\{\phi, h\}\| \\
&\quad =\left\|\sum_{k=1}^{m}(I+t_NA_N)^{-(m-k)}P_NT(t_N)^k\{\phi, h\}\right. \\
&\qquad \left. -(I+t_NA_N)^{-(m-k+1)}P_NT(t_N)^{k-1}\{\phi, h\}\right\| \\
&\quad \leqq \sum_{k=1}^{m}(1-t_N\gamma_N)^{-(m-k)}\|P_NT(kt_N)\{\phi, h\} \\
&\qquad\qquad -(I+t_NA_N)^{-1}P_NT((k-1)t_N)\{\phi, h\}\| \\
&\quad \leqq (1-t_N\gamma_N)^{-m}\sum_{k=1}^{m}\|P_NT((k-1)t_N)\{\phi, h\}-(I+t_NA_N)P_NT(kt_N)\{\phi, h\}\|.
\end{aligned} \tag{2.21}$$

The remainder of the proof is just as in the proof of Theorem 2.1.

Remark 2.1. Let F be continuously Fréchet differentiable from $L^2(-r, 0; R^n)$ to R^n. Let $\{\phi, h\}\in D(A)$, let $\phi''\in L^\infty(-r, 0; R^n)$, let $\phi'(0)=F(\phi)$, and let

$$|F(\pi_1 T(t)\{\phi, h\})-F_N(\pi_1 P_N T(t)\{\phi, h\})|=O(1/\sqrt{N})$$

uniformly on bounded t-intervals. From formulas (2.16) and (2.17) we have that

$\dot{x}(\phi, h)(t)$ is differentiable from $[-r, \infty)$ to R^n and

$$\ddot{x}(\phi, h)(t)=\begin{cases}\phi'' & \text{for a.e. } t\in[-r, 0],\\ F'(x_t(\phi, h))\pi_1 AT(t)\{\phi, h\} & \text{for } t>0.\end{cases}$$

Using (1.4) we see that for $0\leqq s\leqq t$

$$\begin{aligned}&\|AT(t)\{\phi, h\}-AT(s)\{\phi, h\}\|^2\\ &\quad=\int_{-r}^{0}|(\pi_1 AT(t)\{\phi, h\})(\theta)-(\pi_1 AT(s)\{\phi, h\})(\theta)|^2\,d\theta\\ &\qquad+|\pi_2 AT(t)\{\phi, h\}-\pi_2 AT(s)\{\phi, h\}|^2\\ &\quad=\int_{-r}^{0}|\dot{x}(\phi, h)(t+\theta)-\dot{x}(\phi, h)(s+\theta)|^2\,d\theta\\ &\qquad+|F(\pi_1 T(t)\{\phi, h\})-F(\pi_1 T(s)\{\phi, h\})|^2\\ &\quad\leqq \text{const.}\ |t-s|^2+\beta^2|t-s|^2 e^{2\gamma t}\|A\{\phi, h\}\|^2.\end{aligned}$$

Then, from (2.18), (2.19), and (2.20) one may verify that

$$\|(I-t_N A_N)^{[t/t_N]}P_N\{\phi, h\}-P_N T(t)\{\phi, h\}\|=O(1/\sqrt{N}),$$

and, similarly, one may verify that

$$\|(I+t_N A_N)^{[t/t_N]}P_N\{\phi, h\}-P_N T(t)\{\phi, h\}\|=O(1/\sqrt{N}).$$

Remark 2.2. Since A_N is a Lipschitz continuous operator in X_N we have that for $t\geqq 0$ and $\{\phi, h\}\in X_N$,

$$T_N(t)\{\phi, h\}\overset{\text{def}}{=}\lim_{m\to\infty}\left(I-\frac{t}{m}A_N\right)^m\{\phi, h\}=\lim_{m\to\infty}\left(I+\frac{t}{m}A_N\right)^{-m}\{\phi, h\} \tag{2.22}$$

exists and defines a nonlinear semigroup $T_N(t)$, $t\geqq 0$ in X_N. By virtue of Lemma 2.1 or 2.2

$$\|T_N(t)\{\phi, h\}-T_N(t)\{\psi, k\}\|\leqq e^{\gamma_N t}\|\{\phi, h\}-\{\psi, k\}\|.$$

Moreover, we can show that for $t\geqq 0$, $\{\phi, h\}\in X$,

$$\lim_{N\to\infty}\|T_N(t)P_N\{\phi, h\}-P_N T(t)\{\phi, h\}\|=0. \tag{2.23}$$

To see this claim let $t\geqq 0$, $\{\phi, h\}\in X$, $\varepsilon>0$, and choose an increasing sequence of positive integers $\{m_N\}_{N=1}^{\infty}$ such that $t/m_N\leqq r/N$ and

$$\|(I-(t/m_N)A_N)^{m_N}P_N\{\phi, h\}-T_N(t)P_N\{\phi, h\}\|<\varepsilon.$$

Set $t_N=t/m_N$ and appeal to Theorem 2.1 to choose an integer N_1 sufficiently large such that if $N\geqq N_1$, then

$$\|(I-t_N A_N)^{[t/t_N]}P_N\{\phi, h\}-P_N T(t)\{\phi, h\}\|<\varepsilon.$$

Then, (2.23) follows immediately.

Remark 2.3. For computations there is a considerable advantage to the explicit approximation result of Theorem 2.1 as opposed to the implicit approximation result of Theorem 2.2. This advantage is illustrated in Example 2.1, where for

$\{\sum_{i=1}^{N} h_i\chi_i^N, h_0\}$ given in X_N,

$$\left\{\sum_{i=1}^{N} k_i\chi_i^N, k_0\right\} \overset{\text{def}}{=} (I - t_N A_N)\left\{\sum_{i=1}^{N} h_i\chi_i^N, h_0\right\} \tag{2.24}$$

is equivalent to the explicit system of $N+1$ equations in $N+1$ unknowns $k_0, \cdots, k_N$,

$$\begin{aligned} k_0 &= h_0 + t_N g_N\left(\frac{r}{N}\sum_{i=1}^{N} f_N(h_i)\right), \\ k_i &= (1-(t_N N/r))h_i + (t_N N/r)h_{i-1}, \qquad 1 \leqq i \leqq N, \end{aligned} \tag{2.25}$$

and

$$\left\{\sum_{i=1}^{N} k_i\chi_i^N, k_0\right\} \overset{\text{def}}{=} (I + t_N A_N)^{-1}\left\{\sum_{i=1}^{N} h_i\chi_i^N, h_0\right\} \tag{2.26}$$

is equivalent to the implicit system of $N+1$ equations in $N+1$ unknowns $k_0, \cdots, k_N$,

$$\begin{aligned} h_0 &= k_0 - t_N g_N\left(\frac{r}{N}\sum_{i=1}^{N} f_N(k_i)\right), \\ h_i &= (1+(t_N N/r))k_i - (t_N N/r)k_{i-1}, \qquad 1 \leqq i \leqq N. \end{aligned} \tag{2.27}$$

There is an advantage of the implicit scheme of Theorem 2.2 in that $\{t_N\}_{N=1}^{\infty}$ need not decrease to 0 as fast as $\{r/N\}_{N=1}^{\infty}$, so that $[t/t_N]$ might be smaller than $[tN/r]$ in (2.9). Note that in case $t_N = r/N$, (2.25) becomes simply

$$\begin{aligned} k_0 &= h_0 + t_N g_N\left(t_N \sum_{i=1}^{N} f_N(h_i)\right), \\ k_i &= h_{i-1}, \qquad i \leqq i \leqq N. \end{aligned}$$

Thus our explicit method may be viewed as an adaptation of Eulers one step method. The linear equations in (2.27) may be solved explicitly in terms of k_0. Then one finds

$$k_j = \left(\frac{\alpha}{1+\alpha}\right)^j k_0 + \sum_{i=1}^{j} \frac{\alpha^{i-1}}{(1+\alpha)^i} h_{j-i+1}, \qquad \alpha = t_N N/r. \tag{2.28}$$

Substituting these formulas into the first equation in (2.27), our implicit method then may be seen to require the solution of one nonlinear equation in one unknown, namely k_0, at each step and can be computed by the method of successive substitutions for t_N sufficiently small. Formula (2.28) should not be computed as given. Suppose $\alpha = 1$ and let $c = \alpha/(1+\alpha) = .5$. Define $\xi_0 = k_0$ and define $\xi_i = c(\xi_{i-1} + h_i)$, $i = 1, \cdots, N$. Then (2.27) may be written as

$$h_0 = k_0 - t_N g_N\left(\frac{r}{N}\sum_{i=1}^{N} f_N(\xi_i)\right).$$

Thus using this recursion to produce $\{k_i\}$ from k_0 and $\{h_i\}$, only N multiplications are required.

3. The discontinuous case. In this case we study (FDE) with F having the special form

$$F(\phi) = g\left(\int_{-r}^{0} d\eta(\theta)\phi(\theta)\right), \qquad D(F) = C(-r, 0; R^n). \tag{3.1}$$

In (3.1) we require that $g: R^n \to R^n$ is Lipschitz continuous with Lipschitz constant β and $\eta: [-r, 0] \to \mathrm{Lip}(R^n, R^n)$ is of bounded variation from $[-r, 0]$ to $\mathrm{Lip}(R^n, R^n)$ [that is, there exists a number K such that if $\{\theta_i\}_{i=0}^N$ is a subdivision of $[-r, 0]$, then $\sum_{i=1}^N \|\eta(\theta_i)-\eta(\theta_{i-1})\|_{\mathrm{Lip}} \leqq K$, and the total variation of η between $-r$ and 0, denoted by $\int_{-r}^0 |d\eta|$, is the least such number K]. For $\phi \in C(-r, 0; R^n)$, the Stieltjes integral $\int_{-r}^0 d\eta(\theta)\phi(\theta)$ exists in the sense that there exists $x \in R^n$ such that if $\varepsilon > 0$ there exists a subdivision $\{\hat{\theta}_i\}_{i=0}^M$ of $[-r, 0]$ such that if $\{\theta_i\}_{i=0}^N$ is a refinement of $\{\hat{\theta}_i\}_{i=0}^M$, $\theta_{i-1} \leqq \theta_i' \leqq \theta_i$, then $|\sum_{i=1}^N (\eta(\theta_u)-\eta(\theta_{i-1}))\phi(\theta_i')-x| < \varepsilon$. That $\int_{-r}^0 d\eta(\theta)\phi(\theta)$ exists for $\phi \in C(-r, 0; R^n)$ is proved using the following: if $\{\theta_i\}_{i=0}^N$ is a subdivision of $[-r, 0]$, $\{\delta_j\}_{j=0}^M$ is a refinement of $\{\theta_i\}_{i=0}^N$ with $\delta_{k_i} = \theta_i$, $k_0 = 0 < k_1 < \cdots < k_N = M$, then

$$\begin{aligned}
&\left|\sum_{i=1}^N (\eta(\theta_i)-\eta(\theta_{i-1}))\phi(\theta_i') - \sum_{j=1}^M (\eta(\delta_j)-\eta(\delta_{j-1}))\phi(\delta_j')\right| \\
&\quad = \left|\sum_{i=1}^N \sum_{j=k_{i-1}+1}^{k_i} ((\eta(\delta_j)-\eta(\delta_{j-1}))\phi(\theta_i') - (\eta(\delta_j)-\eta(\delta_{j-1}))\phi(\delta_j'))\right| \\
&\quad \leqq \sum_{i=1}^N \sum_{j=k_{i-1}+1}^{k_i} |\eta(\delta_j)-\eta(\delta_{j-1})|_{\mathrm{Lip}} |\phi(\theta_i')-\phi(\delta_j')| \\
&\quad \leqq \sup_{1\leqq i\leqq N} \sup_{k_{i-i}+1\leqq j\leqq k_i} |\phi(\theta_i')-\phi(\delta_j')| \int_{-r}^0 |d\eta|.
\end{aligned}$$

[The Cauchy convergence condition thus holds and this implies the existence of the integral (see [16, 10.1, p. 49]).] Further, $\phi \to \int_{-r}^0 d\eta(\theta)\phi(\theta)$ is Lipschitz continuous form $C(-r, 0; R^n)$ to R^n, since

$$\begin{aligned}
&\left|\sum_{i=1}^N (\eta(\theta_i)-\eta(\theta_{i-1}))\phi(\theta_i') - \sum_{i=1}^N (\eta(\theta_i)-\eta(\theta_{i-1}))\psi(\theta_i')\right| \\
&\quad \leqq \sum_{i=1}^N |(\eta(\theta_i)-\eta(\theta_{i-1}))\phi(\theta_i') - (\eta(\theta_i)-\eta(\theta_{i-1}))\psi(\theta_i')| \\
&\quad \leqq \sum_{i=1}^N |\eta(\theta_i)-\eta(\theta_{i-1})|_{\mathrm{Lip}} |\phi(\theta_i')-\psi(\theta_i')| \\
&\quad \leqq \sup_{-r\leqq\theta\leqq 0} |\phi(\theta)-\psi(\theta)| \sum_{i=1}^N \|\eta(\theta_i)-\eta(\theta_{i-1})\|_{\mathrm{Lip}} \\
&\quad \leqq \sup_{-r\leqq\theta\leqq 0} |\phi(\theta)-\psi(\theta)| \int_{-r}^0 |d\eta|.
\end{aligned}$$

We also require that $\eta(0) = 0$ and $\lim_{\theta\to -r} \tau(\theta) \neq 0$, where $\tau(\theta) \overset{\text{def}}{=} \int_{-r}^\theta |d\eta|$ (the total variation of η between $-r$ and θ). As a mapping from $L^2(-r, 0; R^n)$ to R^n such an F may be only densely defined and discontinuous, but this class of F includes equations of delay type not included in the class of F considered in § 2.

Again we associate with the solutions of (FDE) a semigroup of nonlinear operators [15, Props. 4.1, 4.2]. In this case the semigroup $T(t)$, $t \geqq 0$, is defined in the space $X_\mu \overset{\text{def}}{=} L^2(-r, 0; R^n; \mu) \times R^n$, $d\mu(\theta) \overset{\text{def}}{=} \tau(\theta)\, d\theta$, with norm and inner product

$$\begin{aligned}
&\|\{\phi, h\}\|_\mu^2 = \int_{-r}^0 |\phi(\theta)|^2 \tau(\theta)\, d\theta + |h|^2, \\
&\langle\{\phi, h\}, \{\psi, k\}\rangle_\mu = \int_{-r}^0 (\phi(\theta), \psi(\theta))\tau(\theta)\, d\theta + (h, k)
\end{aligned} \tag{3.2}$$

and $T(t)$, $t \geqq 0$, satisfies (1.1)–(1.5) with $\gamma = \tau(0)(1+\beta^2)/2$. If we do not introduce the weighted norm in (3.2), then the semigroup $T(t)$, $t \geqq 0$, associated with (FDE) will satisfy, in general, the condition that the Lipschitz constant of $T(t)$ is $\leqq Me^{\omega t}$, with $M > 1$. In both the linear and nonlinear case it is very advantageous to have this constant M equal to 1, especially for approximation theory.

In order to show the stability of an approximating scheme for a noncontraction type semigroup in the linear case, one is required to show that the approximating semigroups T_N satisfy a condition of the form $|T_N(t)| \leqq Me^{\omega t}$ uniformly in N [11, p. 502]. In the nonlinear case there is at present no approximation theory available for noncontraction type semigroups comparable to that for contraction type semigroups (general treatments of approximation theory for contraction type semigroups may be found in [4] and [9]). As in § 2 we identify the solutions of (FDE) with $T(t)$, $t \geqq 0$, by means of the formulas (2.2). But for F as in (3.1) we know only that $x(\phi, h)(t)$ satisfies (FDE) provided $\{\phi, h\} \in D(A)$ [15, Prop. 5.8]. In general we must think of the function $x(\phi, h)(t)$ defined in (2.2) as a solution of (FDE) in a generalized sense [15, Prop. 5.12 and Cor. 5.13]. The generator of $T(t)$, $t \geqq 0$, is $A: X_\mu \to X_\mu$, where

$$D(A) = \{\{\phi, h\}: \phi \text{ is absolutely continuous, } \phi' \in L^2(-r, 0; R^n; \mu), \phi(0) = h\},$$
(3.3)
$$A\{\phi, h\} = \{-\phi', -F(\phi)\}.$$

In § 2 the approximations to the solutions of (FDE) were defined in a subspace of X. In this section the approximations will be defined in a space X_N equipped with a norm different from the norm of X_μ. The norm of X_N will arise from a discretized version of the measure μ in the norm of X_μ, and this "weighting" of the norm of X_N will yield the stability of our approximating schemes in an easy and natural way. For each positive integer N define $\tau_j^N = \tau(-rj/N)$, $j = 0, 1, \cdots, N$ and define the Hilbert space X_N by

$$X_N = \left\{\{\phi, h\}: \phi = \sum_{i=1}^{N} h_i \chi_i^N, h_1, \cdots, h_N \in R^n\right\},$$
(3.4)
$$\|\{\phi, h\}\|_N^2 = \frac{r}{N} \sum_{i=1}^{N} |h_i|^2 \tau_{i-1}^N + |h|^2,$$
$$\langle \{\phi, h\}, \{\psi, k\} \rangle_N = \frac{r}{N} \sum_{i=1}^{N} (h_i, k_i) \tau_{i-1}^N + (h, k).$$

Define P_N as in (2.4) except that now P_N maps X_μ into X_N. Since for $\{\phi, h\} \in X_N$

$$\begin{aligned} \|P_N\{\phi, h\}\|_N^2 &= \frac{r}{N} \sum_{i=1}^{N} \left| \frac{N}{r} \int_{-ri/N}^{-r(i-1)/N} \phi(\theta)\, d\theta \right|^2 \tau_{i-1}^N + |h|^2 \\ &\leqq \sum_{i=1}^{N} \left(\int_{-ri/N}^{-r(i-1)/N} |\phi(\theta)|^2\, d\theta \right) \tau_{i-1}^N + |h|^2 \\ &\leqq \frac{\tau(0)}{\lim_{\theta \to -r} \tau(\theta)} \int_{-r}^{0} |\phi(\theta)|^2 \tau(\theta)\, d\theta + |h|^2, \end{aligned}$$

the P_N's are uniformly bounded in N as operators from X_μ to X_N.

For each N let g_N be a Lipschitz continuous operator from R^n to R^n with

Lipschitz constant β_N such that

(3.5) $$\sup_{N\geqq 1} \beta_N < \infty$$

(3.6) $$\lim_{N\to\infty} g_N(h) = g(h) \quad \text{for all } h \in R^n.$$

We note that (3.6) holds uniformly on compact subsets of R^n by virtue of (3.5). Let π_1^N be the projection on X_N defined by $\pi_1^N\{\phi, h\} = \phi$ and let $F_N: \pi_1^N X_N \to R^n$ be defined by

$$F_N\left(\sum_{i=1}^N h_i \chi_i^N\right) = g_N\left(\sum_{i=1}^N \left(\eta\left(\frac{-r(i-1)}{N}\right) - \eta\left(\frac{-ri}{N}\right)\right) h_i\right)$$

for $\sum_{i=1}^N h_i \chi_i^N \in \pi_1^N X_N$.

Example 3.1. Let f, a, $b \in \mathrm{Lip}\,(R, R)$ and let $0 \leqq r_2 \leqq r_1 = r$. Consider the second-order scalar delay equation

$$\ddot{y}(t) = f(a(y(t-r_1)) + b(\dot{y}(t-r_2))), \qquad t \geqq 0$$
$$y_0 = \phi, \quad y(0) = h, \quad \dot{y}_0 = \hat{\phi}, \quad \dot{y}(0) = \hat{h}, \quad \phi, \hat{\phi} \in L^2(-r, 0; R), \quad h, \hat{h} \in R.$$

Define $\eta: [-r, 0] \to \mathrm{Lip}\,(R^2, R^2)$ by $\eta(\theta)(h, \hat{h}) = (0, 0)$ if $\theta = -r_1$, $(0, a(h))$ if $-r_1 < \theta \leqq -r_2$, $(0, a(h) + b(\hat{h}))$ if $r_2 < \theta < 0$, and $(\hat{h}, a(h) + b(\hat{h}))$ if $\theta = 0$. If $(\phi, \hat{\phi}) \in C(-r, 0; R^2)$, then $\int_{-r}^0 d\eta(\theta)(\phi(\theta), \hat{\phi}(\theta)) = (\hat{\phi}(0), \;\; a(\phi(-r_1)) + b(\hat{\phi}(-r_2)))$. Define $g: R^2 \to R^2$ by $g(h, \hat{h}) = (h, f(\hat{h}))$ and define $F: C(-r, 0; R^2) \to R^2$ by $F(\phi, \hat{\phi}) = g(\int_{-r}^0 d\eta(\theta)(\phi(\theta), \hat{\phi}(\theta))) = (\hat{\phi}(0), f(a(\phi(-r_1)) + b(\hat{\phi}(-r_2))))$. Our second-order scalar delay equation may now be formulated as the first-order equation in R^2

$$\dot{x}(t) = F(x_t), \quad x_0 = (\phi, \hat{\phi}), \quad x(0) = (h, \hat{h}),$$

where $x(t): [-r, \infty) \to R^2$ and $(y(t), \dot{y}(t)) \overset{\text{def}}{=} x(t)$. If we let $\pi(h, \hat{h}) = h$, $\hat{\pi}(h, \hat{h}) = \hat{h}$, $f_N \to f$ in $\mathrm{Lip}\,(R, R)$ and define $g_N(h, \hat{h}) = (h, f_N(\hat{h}))$, then

$$F_N\left(\sum_{i=1}^N (h_i, \hat{h}_i)\chi_i^N\right) = \left(\pi \sum_{i=1}^N \left(\eta\left(\frac{-r(i-1)}{N}\right) - \eta(-ri/N)\right)(h_i, \hat{h}_i),\right.$$
$$\left. f_N\left(\hat{\pi} \sum_{i=1}^N \left(\eta\left(\frac{-r(i-1)}{N}\right) - \eta\left(\frac{-ri}{N}\right)\right)(h_i, \hat{h}_i)\right)\right)$$
$$= \left(\hat{h}_1, f_N\left(\hat{\pi} \sum_{i=1}^N \left(\eta\left(\frac{-r(i-1)}{N}\right) - \eta\left(\frac{-ri}{N}\right)\right)(h_i, \hat{h}_i)\right)\right).$$

We need the following facts about the operators F_N:

LEMMA 3.1. *For each N, F_N is continuous (but not uniformly in N). If $\{\phi, h\} \in D(A)$, then*

(3.7) $$\lim_{N\to\infty} F_N(\pi_1^N P_N T(t)\{\phi, h\}) = F(\pi_1 T(t)\{\phi, h\})$$

uniformly in bounded t intervals.

Proof. If $\sum_{i=1}^N h_i\chi_i^N, \sum_{i=1}^N k_i\chi_i^N \in \pi_i^N X_N$, then the first claim follows from

$$\tag{3.8} \begin{aligned} &\left|F_N\left(\sum_{i=1}^N h_i\chi_i^N\right)-F_N\left(\sum_{i=1}^N k_i\chi_i^N\right)\right| \\ &\qquad \leqq \beta_N \sum_{i=1}^N \left|\eta\left(\frac{-r(i-1)}{N}\right)-\eta\left(\frac{-ri}{N}\right)\right|_{\text{Lip}} |h_i-k_i| \\ &\qquad \leqq \beta_N \sum_{i=1}^N \int_{-ri/N}^{-r(i-1)/N} |d\eta|\,|h_i-k_i| \end{aligned}$$

$$\tag{3.8a} \begin{aligned} &\qquad = \beta_N \sum_{i=1}^N (\tau_{i-1}^N-\tau_i^N)|h_i-k_i| \\ &\qquad \leqq \beta_N \sum_{i=1}^N \tau_{i-1}^N |h_i-k_i| \\ &\qquad \leqq \beta_N \left(\frac{N}{r}\right)^{1/2}\left(\sum_{i=1}^N \tau_{i-1}^N\right)^{1/2}\left(\frac{r}{N}\sum_{i=1}^N \tau_{i-1}^N|h_i-k_i|^2\right)^{1/2}. \end{aligned}$$

If $\{\phi, h\}\in D(A)$ and $\psi \overset{\text{def}}{=} \pi_1 T(t)\{\phi, h\}$, then the second claim follows from (1.3), (3.5), (3.6), and the estimates

$$\begin{aligned} &\left|g_N\left(\sum_{i=1}^N\left(\eta\left(\frac{-r(i-1)}{N}\right)-\eta\left(\frac{-ri}{N}\right)\right)\frac{N}{r}\int_{-ri/N}^{-r(i-1)/N}\psi(\theta)\,d\theta\right)\right. \\ &\qquad \left. -g_N\left(\sum_{i=1}^N\left(\eta\left(\frac{-r(i-1)}{N}\right)-\eta\left(\frac{-ri}{N}\right)\right)\psi\left(\frac{-ri}{N}\right)\right)\right| \\ &\leqq \beta_N \sum_{i=1}^N \int_{-ri/N}^{-r(i-1)/N}|d\eta|\left|\frac{N}{r}\int_{-ri/N}^{-r(i-1)/N}\psi(\theta)\,d\theta-\psi(-ri/N)\right| \\ &\leqq \beta_N\tau(0)\frac{N}{r}\sup_{1\leqq i\leqq N}\int_{-ri/N}^{-r(i-1)/N}\left|\psi(\theta)-\psi\left(\frac{-ri}{N}\right)\right|d\theta \\ &\leqq \beta_N\tau(0)\sup_{1\leqq i\leqq N,-ri/N\leqq\theta\leqq -r(i-1)/N}\left|\psi(\theta)-\psi\left(\frac{-ri}{N}\right)\right| \\ &\leqq \beta_N\tau(0)\left(\lim_{\theta\to -r}\tau(\theta)\right)^{-1/2}\left(\frac{r}{N}\right)^{1/2}\left(\int_{-r}^0|\dot\psi(s)|^2\tau(s)\,ds\right)^{1/2}, \end{aligned}$$

and

$$\begin{aligned} &\left|g_N\left(\sum_{i=1}^N\left(\eta\left(\frac{-r(i-1)}{N}\right)-\eta\left(\frac{-ri}{N}\right)\right)\psi\left(\frac{-ri}{N}\right)\right)-g_N\left(\int_{-r}^0 d\eta(\theta)\psi(\theta)\right)\right| \\ &= \left|g_N\left(\sum_{i=1}^N\int_{-ri/N}^{-r(i-1)/N}d\eta(\theta)\psi\left(\frac{-ri}{N}\right)\right)-g_N\left(\sum_{i=1}^N\int_{-ri/N}^{-r(i-1)/N}d\eta(\theta)\psi(\theta)\right)\right| \\ &\leqq \beta_N\sum_{i=1}^N\int_{-ri/N}^{-r(i-1)/N}|d\eta|\sup_{-ri/N\leqq\theta\leqq -r(i-1)/N}\left|\psi\left(\frac{-ri}{N}\right)-\psi(\theta)\right| \\ &\leqq \beta_N\tau(0)\left(\lim_{\theta\to -r}\tau(\theta)\right)^{-1/2}\left(\frac{r}{N}\right)^{1/2}\left(\int_{-r}^0|\dot\psi(s)|^2\tau(s)\,ds\right)^{1/2}. \end{aligned}$$

For each N we define a discrete operator A_N approximating A by $A_N\colon X_N \to X_N$,

$$\text{(3.9)}\qquad \begin{aligned} &D(A_N)=X_N \quad \text{and} \quad \text{for} \quad \{\phi, h\}=\left\{\sum_{i=1}^{N} h_i\chi_i^N, h_0\right\} \\ &A_N\{\phi, h\}=\left\{-\sum_{i=1}^{N} \frac{N}{r}(h_{i-1}-h_i)\chi_i^N, -F_N(\phi)\right\}. \end{aligned}$$

THEOREM 3.1. *If* $\{t_N\}_{N=1}^{\infty}$ *is a decreasing sequence of positive numbers such that* $t_N \leqq r/N$, $\{\phi, h\}\in X$, *and* $t \geqq 0$, *then*

$$\text{(3.10)}\qquad \lim_{N\to\infty} \|(I-t_NA_N)^{[t/t_N]}P_N\{\phi, h\}-P_NT(t)\{\phi, h\}\|_N=0.$$

THEOREM 3.2. *If* $\{t_N\}_{N=1}^{\infty}$ *is a decreasing sequence of positive numbers converging to* 0, $\{\phi, h\}\in X$, *and* $t\geqq 0$, *then*

$$\text{(3.11)}\qquad \lim_{N\to\infty} \|(I+t_NA_N)^{-[t/t_N]}P_N\{\phi, h\}-P_NT(t)\{\phi, h\}\|_N=0.$$

Before proving Theorems 3.1 and 3.2 we establish the stability of the approximating schemes in the following lemmas.

LEMMA 3.2. *For each* N *let* $\gamma_N=\tau(0)(1+\beta_N^2)/2$. *If* $0<\lambda\leqq r/N$ *and* $\{\phi, h\}$, $\{\psi, k\}\in X_N$, *then*

$$\text{(3.12)}\qquad \|(I-\lambda A_N)\{\phi, h\}-(I-\lambda A_N)\{\psi, k\}\|\leqq (1+\lambda\gamma_N)\|\{\phi, h\}-\{\psi, k\}\|.$$

Proof. Let $\{\phi, h\}=\{\sum_{i=1}^N h_i\chi_i^N, h_0\}$, $\{\psi, k\}=\{\sum_{i=1}^N k_i\chi_i^N, k_0\}$, and $u_i=h_i-k_i$, $i=0, 1, \cdots, N$. As in the proof of Lemma 2.1 the left side of (3.12) is $\leqq$

$$\left(1-\lambda\left(\frac{N}{r}\right)\right)\|\{\phi, h\}-\{\psi, k\}\|+\frac{\lambda N}{r}\left\|\left\{\sum_{i=1}^N u_{i-1}\chi_i^N, u_0+\frac{r}{N}(F_N(\phi)-F_N(\psi))\right\}\right\|$$

$$=\left(1-\lambda\left(\frac{N}{r}\right)\right)\|\{\phi, h\}-\{\psi, k\}\|+\left(\frac{\lambda N}{r}\right)\left\|\left(I-\frac{r}{N}A_N\right)\{\phi, h\}-\left(I-\frac{r}{N}A_N\right)\{\psi, k\}\right\|.$$

Observe that by the Cauchy–Schwarz inequality

$$\text{(3.13)}\qquad \begin{aligned} \left(\sum_{i=1}^N (\tau_{i-1}^N-\tau_i^N)|u_i|\right)^2 &\leqq \left(\sum_{i=1}^N (\tau_{i-1}^N-\tau_i^N)|u_i|^2\right)\left(\sum_{i=1}^N \tau_{i-1}^N-\tau_i^N\right) \\ &=\tau(0)\sum_{i=1}^N |u_i|^2(\tau_{i-1}^N-\tau_i^N). \end{aligned}$$

Then, (3.12) follows using (3.8a), (3.13), and

$$\begin{aligned} &\left\|\left(I-\frac{r}{N}A_N\right)\{\phi, h\}-\left(I-\frac{r}{N}A_N\right)\{\psi, k\}\right\|^2 \\ &=\frac{r}{N}\sum_{i=1}^N |u_{i-1}|^2\tau_{i-1}^N+|u_0|^2+\frac{2r}{N}(u_0, F_N(\phi)-F_N(\psi))+\left(\frac{r}{N}\right)^2|F_N(\phi)-F_N(\psi)|^2 \\ &\leqq\frac{r}{N}\sum_{i=1}^N |u_{i-1}|^2\tau_{i-1}^N+|u_0|^2+\frac{2r}{N}|u_0|\beta_N\sum_{i=1}^N (\tau_{i-1}^N-\tau_i^N)|u_i| \\ &\quad+\left(\frac{r}{N}\right)^2\beta_N^2\left(\sum_{i=1}^N (\tau_{i-1}^N-\tau_i^N)|u_i|\right)^2 \end{aligned}$$

$$\leqq \frac{r}{N}\sum_{i=1}^{N}|u_{i-1}|^2\tau_{i-1}^N+|u_0|^2+\frac{r}{N}(|u_0|\beta_N\tau(0)^{1/2})^2$$

$$+\frac{r}{N}\left(\frac{1}{\tau(0)^{1/2}}\sum_{i=1}^{N}(\tau_{i-1}^N-\tau_i^N)|u_i|\right)^2+\left(\frac{r}{N}\right)^2\beta_N^2\left(\sum_{i=1}^{N}(\tau_{i-1}^N-\tau_i^N)|u_i|\right)^2$$

$$\leqq \frac{r}{N}\sum_{i=1}^{N}|u_{i-1}|^2\tau_{i-1}^N+|u_0|^2+\frac{r}{N}|u_0|^2\beta_N^2\tau(0)$$

$$+\frac{r}{N}\sum_{i=1}^{N}(\tau_{i-1}^N-\tau_i^N)|u_i|^2+\left(\frac{r}{N}\right)^2\beta_N^2\tau(0)\sum_{i=1}^{N}(\tau_{i-1}^N-\tau_i^N)|u_i|^2$$

$$=\left(1+\left(\frac{2r}{N}\right)\gamma_N\right)|u_0|^2+\frac{r}{N}\sum_{i=1}^{N}|u_i|^2\tau_{i-1}^N-\frac{r}{N}|u_N|^2\tau_N^N$$

$$+\left(\frac{r}{N}\right)^2\beta_N^2\tau(0)\sum_{i=1}^{N}(\tau_{i-1}^N-\tau_i^N)|u_i|^2$$

$$\leqq\left(1+\frac{2r}{N}\gamma_N\right)|u_0|^2+\frac{r}{N}\sum_{i=1}^{N}|u_i|^2\tau_{i-1}^N$$

$$+\left(\frac{r}{N}\right)^2 2\gamma_N\sum_{i=1}^{N}|u_i|^2\tau_{i-1}^N$$

$$\leqq\left(1+\frac{r}{N}\gamma_N\right)^2\|\{\phi,h\}-\{\psi,k\}\|^2.$$

LEMMA 3.3. *For each N let $\gamma_N=\tau(0)(1+\beta_N^2)/2$. Then $A_N+\gamma_N I$ is accretive in X_N. Consequently, for $0<\lambda<1/\gamma_N$, $(I+\lambda A_N)^{-1}$ exists, is everywhere defined on X_N, and satisfies*

$$|(I+\lambda A_N)^{-1}|_{\text{Lip}}\leqq(1+\lambda\gamma_N)^{-1}. \tag{3.14}$$

Proof. Let $\{\phi,h\}$, $\{\psi,k\}\in X_N$, $\phi=\sum_{i=1}^N h_i\chi_i^N$, $h=h_0$, $\psi=\sum_{i=1}^N k_i\chi_i^N$, $k=k_0$, and $u_i=h_i-k_i$, $i=0,1,\cdots,N$. Using (3.8a) we obtain

$$\langle A_N\{\phi,h\}-A_N\{\psi,k\},\{\phi,h\}-\{\psi,k\}\rangle_N$$

$$=\frac{r}{N}\sum_{i=1}^{N}-\frac{N}{r}(u_{i-1}-u_i,u_i)\tau_{i-1}^N-(F_N(\phi)-F_N(\psi),u_0)$$

$$\geqq\sum_{i=1}^{N}-(u_{i-1}-u_i,u_i)\tau_{i-1}^N-\beta_N\sum_{i=1}^{N}(\tau_{i-1}^N-\tau_i^N)|u_i|\,|u_0|$$

$$=\sum_{i=1}^{N}\sum_{k=1}^{i}-(u_{k-1}-u_k,u_k)(\tau_{i-1}^N-\tau_i^N)-\beta_N\sum_{i=1}^{N}(\tau_{i-1}^N-\tau_i^N)|u_i|\,|u_0| \tag{3.15}$$

$$=\sum_{i=1}^{N}\left(\sum_{k=1}^{i}-(u_{k-1}-u_k,u_k)-\beta_N|u_i|\,|u_0|\right)(\tau_{i-1}^N-\tau_i^N)$$

$$\geqq\sum_{i=1}^{N}\left(\sum_{k=1}^{i}\tfrac12(|u_{k-1}|-|u_k|)^2+\tfrac12|u_i|^2\right.$$

$$\left.+(\beta_N^2/2)|u_0|^2-\beta_N|u_i|\,|u_0|-((1+\beta_N^2)/2)|u_0|^2\right)(\tau_{i-1}^N-\tau_i^N)$$

We now use the fact that for all $x, y \geqq 0$, $\frac{1}{2}x^2+(\beta_N^2/2)y^2-\beta_N xy \geqq 0$ to conclude that (3.15)$\geqq -((1+\beta_N^2)/2)|u_0|^2\tau(0)$. This proves the accretiveness of $A_N+\gamma_N I$. The second statement of the lemma follows from the accretiveness of $A_N+\gamma_N I$ and the continuity of A_N (see [14]).

LEMMA 3.4. *If* $\{\phi, h\}\in D(A)$ *and* $t\geqq 0$, *then*

$$\begin{aligned}
&\|P_N A T(t)\{\phi, h\}-A_N P_N T(t)\{\phi, h\}\|_N^2 \\
&\leqq \frac{\tau(0)}{\lim_{\theta\to -r}\tau(\theta)}\left(\int_{-r/N}^{0} |(\pi_1^N A T(t)\{\phi, h\})(\theta)|^2\tau(\theta)\,d\theta\right. \\
&\qquad + \frac{1}{3}\left(\frac{\tau(0)}{\lim_{\theta\to -r}\tau(\theta)}\right)\int_{-r}^{-r/N}\left|(\pi_1^N A T(t)\{\phi, h\})(\theta)\right. \\
&\qquad\qquad \left.\left. -(\pi_1^N A T(t)\{\phi, h\})\left(\theta+\frac{r}{N}\right)\right|^2\tau(\theta)\,d\theta\right) \\
&\qquad + |F(\pi_1 T(t)\{\phi, h\})-F_N(\pi_1^N P_N T(t)\{\phi, h\})|^2.
\end{aligned} \tag{3.16}$$

Proof. The proof is very similar to the proof of Lemma 2.3.

LEMMA 3.5. *If* $\{\phi, h\}\in D(A)$, *then*

(3.17) $t\to AT(t)\{\phi, h\}$ *is continuous from* $[0, \infty)$ *to* X;

(3.18) *for* $t\geqq 0$, $|u|<\delta$,

$$\|T(t+u)\{\phi, h\}-T(t)\{\phi, h\}+\mu A T(t)\{\phi, h\}\|_\mu \leqq \delta \max_{t-\delta\leqq s\leqq t+\delta}\|AT(s)\{\phi, h\}-AT(t)\{\phi, h\}\|_\mu;$$

(3.19) $AT(t)\{\phi, h\}=\{(x_t(\phi, h))', \dot{x}(\phi, h)(t)\}$;

(3.20) $\dot{x}(\phi, h)(t)=\phi'(t)$ *for a.e.* $t\in[-r, 0]$ *and* $F(x_t(\phi, h))$ *for* $t\geqq 0$.

Proof. (3.19) and (3.20) follow from (1.3), (2.2), (3.3) and Proposition 5.8 of [15]. Then, (3.17) is established just as in the proof of (2.14) except that we use the fact that F is continuous from $C(-r, 0; R^n)$ to R^n. Statement (3.18) is proved exactly as (2.15) in Lemma 2.4.

Theorems 3.1 and 3.2 are now proved just as Theorems 2.1 and 2.2 were proved except that we use Lemmas 3.1, 3.2, 3.3, 3.4, and 3.5.

Remark 3.1. Let F be continuously Fréchet differentiable from $C(-r, 0; R^n)$ to R^n. Let $\{\phi, h\}\in D(A)$, let $\phi''\in L^\infty(-r, 0; R^n)$, let $\phi'(0)=F(\phi)$, and let

$$|F(\pi_1 T(t)\{\phi, h\})-F_N(\pi_1^N P_N T(t)\{\phi, h\})|=O(1/\sqrt{N})$$

uniformly on bounded t-intervals. Using formulas (3.19) and (3.20) we can establish that the convergence in (3.10) and (3.11) is $O(1/\sqrt{N})$ just as in (2.20) and (2.21) of Remark 2.1.

Remark 3.2. The comments made in Remark 2.2 for the continuous case apply also to the discontinuous case. That is, (2.23) holds for the discontinuous case, where $T_N(t)$, $t\geqq 0$, is defined as in (2.22). The comments made in Remark 2.3 concerning the relative merits of the explicit and implicit schemes also apply for the discontinuous case. For Example 3.1, if $\{\sum_{i=1}^N (h_i, \hat{h}_i)\chi_i^N, (h_0, \hat{h}_0)\}$ is in X_N, then

$$\begin{aligned}\left\{\sum_{i=1}^N (k_i, \hat{k}_i)\chi_i^N, (k_0, \hat{k}_0)\right\} &\overset{\text{def}}{=} (I - t_N A_N)\left\{\sum_{i=1}^N (h_i, \hat{h}_i)\chi_i^N, (h_0, \hat{h}_0)\right\} \\ &= \left\{\sum_{i=1}^N \left(\left(1 - t_N\left(\frac{N}{r}\right)\right)(h_i, \hat{h}_i) + \frac{t_N N}{r}(h_{i-1}, \hat{h}_{i-1})\right)\chi_i^N, (h_0, \hat{h}_0)\right. \\ &\qquad \left. + t_N F_N\left(\sum_{i=1}^N (h_i, \hat{h}_i)\chi_i^N\right)\right\}\end{aligned} \tag{3.21}$$

is equivalent to the explicit system of $2(N+1)$ equations in $2(N+1)$ unknowns, $k_0, \hat{k}_0, \cdots, k_N, \hat{k}_N$, given by

$$\begin{aligned} k_0 &= h_0 + t_N \hat{h}_1, \\ \hat{k}_0 &= \hat{h}_0 + t_N f_N\left(\hat{\pi} \sum_{i=1}^N \left(\eta\left(\frac{-r(i-1)}{N}\right) - \eta\left(\frac{-ri}{N}\right)\right)(h_i, \hat{h}_i)\right), \\ k_i &= (1 - t_N(N/r))h_i + (t_N N/r)h_{i-1}, \qquad i = 1, \cdots, N, \\ \hat{k}_i &= (1 - t_N(N/r))\hat{h}_i + (t_N N/r)\hat{h}_{i-1}, \qquad i = 1, \cdots, N. \end{aligned} \tag{3.22}$$

Note that with $t_N N/r = 1$, we have simply $k_i = h_{i-1}, i = 1, \cdots, N$ and $\hat{k}_i = \hat{h}_{i-1}$, $i = 1, \cdots, N$. Hence apart from shifting indices, (3.22) represents two equations. The equation

$$\left\{\sum_{i=1}^N (k_i, \hat{k}_i)\chi_i^N, (k_0, \hat{k}_0)\right\} \overset{\text{def}}{=} (I + t_N A_N)^{-1}\left\{\sum_{i=1}^N (h_i, \hat{h}_i)\chi_i^N, (h_0, \hat{h}_0)\right\} \tag{3.23}$$

is equivalent to the implicit system of $2(N+1)$ equations in $2(N+1)$ unknowns $k_0, \hat{k}_0, \cdots, k_N, \hat{k}_N$, given by

$$\begin{aligned} h_0 &= k_0 - t_N \hat{k}_1, \\ \hat{h}_0 &= \hat{k}_0 - t_N f_N\left(\hat{\pi} \sum_{i=1}^N \left(\eta\left(\frac{-r(i-1)}{N}\right) - \eta\left(\frac{-ri}{N}\right)\right)(k_i, \hat{k}_i)\right), \\ h_i &= (1 + t_N(N/r))k_i - (t_N N/r)k_{i-1}, \qquad i = 1, \cdots, N, \\ \hat{h}_i &= (1 + t_N(N/r))\hat{k}_i - (t_N N/r)\hat{k}_{i-1}, \qquad 1, \cdots, N. \end{aligned} \tag{3.24}$$

As before, the last two sets of equations in (3.24) may be explicitly solved for k_i and $\hat{k}_i$ in terms of $\{h_i\}$ and k_0 and $\{\hat{h}_i\}$ and $\hat{k}_0$ respectively. Thus we have that (3.24) represents two implicit equations in the two unknowns k_0 and $\hat{k}_0$. Once k_0 and $\hat{k}_0$ are determined, then k_i, $\hat{k}_i$ follow directly from the simple recursion formulas in (3.24).

4. Examples. We applied both the implicit and the explicit methods numerically to the problem

$$\dot{x}(t) = \int_{t-1}^t \cos(x(s))\, ds, \qquad t \geqq 0, \tag{4.1}$$

with

$$\{x_0, x(0)\} = \{1, 1\},$$

i.e. $x(s) = 1$ on $[-1, 0]$. This problem satisfies the hypotheses for our continuous case. In the notation of Example 2.1, $f(x) = \cos(x)$ and $g = I$. Since $\cos x$ can be easily evaluated, we choose $f_N = f$ and $g_N = I$ for all N. Then $F_N(\sum_{i=1}^N h_i \chi_i^N) = (r/N)\sum_{i=1}^N \cos(h_i)$. With $t_N = r/N$, the explicit method becomes, from (2.25),

$$k_0 = h_0 + t_N^2 \sum_{i=1}^N \cos(h_i), \qquad (4.2)$$

$$k_i = h_{i-1}, \qquad 1 \leqq i \leqq N.$$

From (2.27), the implicit method becomes for $t_N = r/N$

$$k_0 = h_0 + t_N^2 \sum_{i=1}^N \cos\left(.5^i k_0 + \sum_{s=1}^{j} .5^s h_{j-s+1}\right), \qquad (4.3a)$$

$$k_i = .5^i k_0 + \sum_{s=1}^{i} .5^s h_{i-s+1}, \qquad i = 1, \dots, N. \qquad (4.3b)$$

Thus once (4.3a) is solved for k_0, (4.3b) explicitly gives the remaining k_i. One can always simply choose $g_N = g$ and $f_N = f$ for all N if f and g are easily computable functions. Note that if $t_N < 1$, then (4.3a) is solvable by the method of successive substitutions. Applying the two methods, we obtained the numerical results for the values of N indicated in Table 1. Note that the error of the implicit method for this example is smaller than that of the explicit method, which somewhat compensates for the additional computations required. For the values of N given, at each step h_0 was used as an initial guess in (4.3a) for k_0. Only one iteration was required then to achieve six significant digits of k_0. The solution has value approximately equal to 1.6007 at $t = 2$ (as is seen by taking N very large).

TABLE 1

N	Abs. Error for Exp. Meth. at $t=1$	Abs. Error for Imp. Meth. at $t=1$	Abs. Error for Exp. Meth. at $t=2$	Abs. Error for Imp. Meth. at $t=2$
10	.0208	.0145	.0333	.0025
20	.0108	.0076	.0155	.0017
40	.0056	.0038	.0072	.0007

We reran this same example, but made the initial function $x(t) = 0$, $-1 \leqq t \leqq 0$. The solution reached approximately 1.528 at $t = 2$. For $N = 40$, the error for the explicit method at $t = 2$ was .013; the error for the implicit method was .001.

Although in the above examples the implicit method is more accurate, the extra computing it requires does not in general justify its choice over the explicit scheme. However, in some important cases, the implicit method would be preferred. In order for the explicit method to converge, the stability condition $t_N \leqq r/N$ must be met. This condition is necessary as we have seen in some other numerical examples. On the other hand, the implicit scheme is unconditionally stable. Thus if a problem has a very small delay, r, or if N must be taken very large so that $P_N\{\phi, h\}$ is an accurate approximation of $\{\phi, h\}$, or if information about the solution for large t is required, the implicit method may then be preferred. We reran the first example with $N = 50$ but with $t_N = 1/10$, i.e. $t_N = 5 \cdot r/N$. The explicit method had an error at $t = 2$ of .0094. The solution to the first example approaches $\pi/2$ as t becomes large. For these last computations with the implicit method ($N = 50$, $t_N = 1/10$), the difference between the computed solution and $\pi/2$ for $t = 10$ was $.45 \cdot 10^{-5}$.

The implicit scheme repeatedly smooths the approximate solution through, for example, (4.3b). This averaging can increase the error as compared to the explicit scheme, particularly for highly oscillatory initial data. We reran the first example but with the initial data $x(t)=1$ on $[-1, -.8)\cup[-.6, -.4]\cup[-.2, 0]$ and $x(t)=-1$ elsewhere on $[-1, 0]$. We point out here that for this special initial data (piecewise constant) $p_N\{\phi, h\}=\{\phi, h\}$ and so no error is introduced initially. Our explicit method for this example can then be viewed as a variant of Euler's method and $O(1/N)$ convergence can be established. Table 2 contains the results we obtained for this problem for several values of N. Note that the convergence of the explicit scheme in this case appears to be $O(1/N)$ and the convergence of the implicit scheme appears to be $O(/1/\sqrt{N})$.

TABLE 2

N	Abs. Error of Exp. Meth. at $t=2$	Abs. Error of Imp. Meth. at $t=2$
10	.0333	.0733
20	.0156	.0485
40	.0071	.0341
80	.0036	.0248

Indeed, the implicit method actually changes the initial data during the first N steps if $t_N=r/N$, which of course introduces error. Thus it would seem computationally advantageous, particularly with an oscillatory initial function, to not use (4.3b) for the first N steps with the implicit method but to retain (4.3a) to define k_0. We did this on the previous example and for $N=40$ obtained an error at $t=2$ of .0059. We made similar runs on other equations and observe essentially the same numerical results.

Acknowledgment. The authors gratefully acknowledge the helpful suggestions of the referee.

REFERENCES

[1] H. T. BANKS AND J. A. BURNS, *An abstract framework for approximate solutions to optimal control problems governed by hereditary systems*, Proc. International Conf. on Diff. Eqs. (Univ. of Southern California, Los Angeles, 1974), Academic Press, New York, 1975, pp. 10–25.
[2] H. T. BANKS AND A. MANITIUS, *Projection series for retarded functional differential equations with applications to optimal control problems*, J. Differential Equations, to appear.
[3] J. G. BORISOVIC AND A. S. TURABIN, *On the Cauchy problem for linear nonhomogeneous differential equations with retarded argument*, Soviet Math. Dokl., 10 (1969), pp. 401–405.
[4] H. BREZIS AND A. PAZY, *Convergence and approximation of semigroups of nonlinear operators in Banach spaces*, J. Functional Anal., 9 (1972), pp. 63–74.
[5] B. D. COLEMAN AND V. J. MIZEL, *Norms and semigroups in the theory of fading memory*, Arch. Rational Mech. Anal., 23 (1966), pp. 87–123.
[6] B. D. COLEMAN AND W. NOLL, *An approximation theorem for functionals with applications in continuum mechanics*, Ibid., 6 (1960), pp. 355–370.
[7] M. G. CRANDALL AND A. PAZY, *Semigroups of nonlinear contractions and dissipative sets*, J. Functional Anal., 3 (1969), pp. 376–418.
[8] M. C. DELFOUR AND L. K. MITTER, *Hereditary differential systems with constant delays. I. General case*, J. Differential Equations, 12 (1972), pp. 213–235.
[9] J. GOLDSTEIN, *Approximation of nonlinear semigroups and evolution equations*, J. Math. Soc. Japan, 24 (1972), pp. 558–573.
[10] J. HALE, *Functional Differential Equations*, Applied Mathematics Series, vol. 3, Springer–Verlag, New York, 1971.

[11] T. KATO, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1966.
[12] R. B. VINTNER, *On the evolution of the state of linear differential delay equations in M^2: properties of the generator*, Electronic Systems Lab. Rep. R 541, Mass. Inst. of Tech., Cambridge, 1974.
[13] G. F. WEBB, *Autonomous nonlinear functional differential equations and nonlinear semigroups*, J. Math. Anal. Appl., 46 (1974), pp. 1–12.
[14] ———, *Continuous nonlinear perturbations of linear accretive operators in Banach spaces*, J. Functional Anal., 10 (1972), pp. 191–203.
[15] ———, *Functional differential equations and nonlinear semigroups in L^p-spaces*, J. Differential Equations, 20 (1976), pp. 71–89.
[16] T. H. HILDEBRANDT, *Introduction to the Theory of Integration*, Pure and Applied Mathematics Series, vol. 13, Academic Press, New York, 1963.

SIAM J. MATH. ANAL.
Vol. 9, No. 6, December 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0015 $01.00/0

MONOTONIC PROPERTIES OF ANALYTIC FUNCTIONS*

A. D. RAWLINS† AND J. D. MORGAN‡

Abstract. A theorem is proved which enables one to obtain monotonic properties of the real part, imaginary part, modulus and phase of an arbitrary analytic function in the complex plane. The monotonic properties are established from the behavior of the analytic function and its derivative on the boundary of the domain in which the monotonic properties are required. As an application some monotonic results are derived for the Bessel functions $J_\nu(z)$ and $H_\nu^{(1)}(z)$, where z is complex and ν is real and positive. The theorem and a corollary can be used to obtain monotonic properties of many other special functions of mathematical physics.

1. Introduction. In the present work a technique is derived enabling one to obtain monotonic properties of the real and imaginary part, (or modulus and phase) of an arbitrary analytic function in the complex plane. The technique requires only a knowledge of the behavior of the derivative (or logarithmic derivative) of the analytic function on the boundary of the domain over which the monotonic properties are required. The method for obtaining these monotonic properties comes from the main theorem (Theorem 1) given below. The proof of Theorem 1 uses the derivative (or logarithmic derivative) of a regular analytic function, the Cauchy–Riemann equations, and the maximum modulus theorem. The crux of the proof relies on the well-known fact that the real and imaginary parts of a function that is regular and analytic in a closed domain, take their extreme values on the boundary of the domain. It will be seen that this theorem offers a powerful tool for determining monotonic properties of a complex function.

As an application of Theorem 1 some monotonic results for the Bessel functions $J_\nu(z)$ and $H_\nu^{(1)}(z)$ are obtained. These monotonic properties yield new Bessel function inequalities, particularly lower bound inequalities.

Throughout the paper we shall assume $z=x(\xi,\eta)+iy(\xi,\eta)$, where x, y, ξ and η are real. The transformation between x, y and the general coordinates ξ, η is assumed to exist and be one to one. A sufficient condition for this is that the Jacobian does not vanish, i.e. $\partial(\xi,\eta)/\partial(x,y)\neq 0$. For a complex function $F(z)$ we note that

$$F(z)=|F(z)|\exp[i\operatorname{Arg}F(z)]=\operatorname{Re}F(z)+i\operatorname{Im}F(z)$$

where $\operatorname{Arg} z$ denotes the principal value of the argument of $F(z)$, given by $-\pi<\operatorname{Arg} z\leqq\pi$.

In the application to the Bessel functions, $g_\nu(z)$ will denote the ratio $J'_\nu(z)/J_\nu(z)$ and $h_\nu(z)$ the ratio $H_\nu^{(1)\prime}(z)/H_\nu^{(1)}(z)$. The order ν is assumed to be real and greater than or equal to zero. Where a branch cut exists in the z-plane it is assumed to be the principal cut along the negative real axis. It should be noticed that although $J_\nu(z)$ has a branch point at the origin for noninteger ν, the function $g_\nu(z)$ has no branch points. Furthermore $g_\nu(z)$ is an odd function of z. Finally we note that similar monotonic properties for $H_\nu^{(2)}(z)$ follow from the monotonic properties for $H_\nu^{(1)}(z)$ via the relationship $\overline{H_\nu^{(2)}(z)}=H_\nu^{(1)}(\bar z)$.

2. Monotonic properties of $F(z)$. In this section we shall prove a general theorem which gives monotonic properties of a regular analytic function. We will

* Received by the editors February 19, 1976, and in revised form May 9, 1977.

† Department of Mathematics, The University, Dundee, Scotland. The work of this author was supported in part by an S.R.C. Fellowship.

‡ Royal Aircraft Establishment, Aero Department, Farnborough, Hants.

require the following Lemma 1 which shows how to construct in certain circumstances, a regular analytic function from a given function that is not regular and analytic.

LEMMA 1. *Let us suppose that a simply connected complex domain D with boundary ∂D is such that if $z \in D \cup \partial D$ the complex function $G(z) = u(x, y) + iv(x, y)$ is not a regular analytic function, but that $|G(z)| \neq 0$ and* Arg $G(z)$ *is harmonic. Then the composite complex function $P(x, y)G(z)$, (where $P(x, y)$ is a real function of x and y) is a regular analytic function when $z \in D \cup \partial D$ if*

$$P(x, y) = \frac{1}{|G(z)|} \exp\left\{-\int^y \frac{\partial}{\partial x}\{\text{Arg } G(z)\}\, dy\right\} \tag{1}$$
$$\equiv \frac{1}{|G(z)|} \exp\left\{\int^x \frac{\partial}{\partial y}\{\text{Arg } G(z)\}\, dx\right\}.$$

Proof. We may write $P(x, y)G(z)$ in the form

$$P(x, y)G(z) = P(x, y)|G(z)| \exp[i \text{ Arg } G(z)], \tag{2}$$

and since $P(x, y)G(z)$ is an analytic function of z, so also is

$$\text{Ln}[P(x, y)G(z)] = \text{Ln}[P(x, y)|G(z)|] + i \text{ Arg } G(z) \tag{3}$$

where we have assumed without loss of generality that $P(x, y) > 0$.

An application of the Cauchy–Riemann equations to (3) gives

$$\frac{\partial}{\partial x}(\text{Ln}[P|G|]) = \frac{\partial}{\partial y} \text{Arg } G \tag{4}$$

$$\frac{\partial}{\partial y}(\text{Ln}[P|G|]) = -\frac{\partial}{\partial x} \text{Arg } G. \tag{5}$$

Provided the consistency relationship $\nabla^2 \text{ Arg } G(z) = 0$ holds equations (4) and (5) can be integrated to give

$$P(x, y)|G(z)| = \exp\left[\int^x \frac{\partial}{\partial y} \text{Arg } G\, dx\right] = \exp\left[-\int^y \frac{\partial}{\partial x} \text{Arg } G\, dy\right]. \tag{6}$$

We are now ready to prove:

THEOREM 1. *Let us suppose that a simply connected complex domain D with boundary ∂D is such that if $z \in D \cup \partial D$ the transformation $z = x(\xi, \eta) + iy(\xi, \eta)$ satisfies the conditions*

$$\frac{\partial \eta}{\partial x} \cdot \frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial y} \cdot \frac{\partial \xi}{\partial y} = 0, \qquad J(z) \equiv \frac{\partial(\xi, \eta)}{\partial(x, y)} = \frac{\partial \xi}{\partial x} \cdot \frac{\partial \eta}{\partial y} - \frac{\partial \xi}{\partial y} \cdot \frac{\partial \eta}{\partial x} \neq 0; \tag{7}$$

$$\nabla^2\{\text{Arg } G(z)\} = 0 \quad \textit{where} \quad G(z) = \frac{\partial \xi}{\partial x} + i\frac{\partial \xi}{\partial y}. \tag{8}$$

If, also, $F(z)$ is a regular analytic function of z for $z \in D \cup \partial D$ then:

(i) Re $F(z)$ *is a monotonic function of ξ, and* Im $F(z)$ *is a monotonic function of η in $D \cup \partial D$ if*

$$\text{Re}\{F'(z)P(x, y)G(z)\} \neq 0, \qquad z \in \partial D; \tag{9}$$

(ii) Re $F(z)$ *is a monotonic function of* η *and* Im $F(z)$ *is a monotonic function of* ξ *in* $D \cup \partial D$ *if*

$$\text{Im}\{F'(z)P(x, y)G(z)\} \neq 0, \qquad z \in \partial D, \tag{10}$$

where in (9) *and* (10) $P(x, y)$ *is given by*

$$P(x, y) = \left\{\left(\frac{\partial \xi}{\partial x}\right)^2 + \left(\frac{\partial \xi}{\partial y}\right)^2\right\}^{-1/2} \exp\left\{-\int^y \frac{\partial}{\partial x} \tan^{-1}\left(\frac{\partial \xi/\partial y}{\partial \xi/\partial x}\right) dy\right\}. \tag{11}$$

Proof. Since $F(z)$ is a regular analytic function of z in $D \cup \partial D$ the Cauchy–Riemann equations give

$$\frac{d}{dz} F(z) = \left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \text{Re } F(z) = \left(\frac{\partial}{\partial y} + i\frac{\partial}{\partial x}\right) \text{Im } F(z).$$

Written in terms of the general coordinates ξ and η these equations become

$$F'(z) = \left\{\overline{G(z)}\frac{\partial}{\partial \xi} + H(z)\frac{\partial}{\partial \eta}\right\} \text{Re } F(z), \tag{12}$$

$$= \left\{i\overline{G(z)}\frac{\partial}{\partial \xi} + iH(z)\frac{\partial}{\partial \eta}\right\} \text{Im } F(z), \tag{13}$$

where

$$H(z) = \frac{\partial \eta}{\partial x} - i\frac{\partial \eta}{\partial y}.$$

Using Lemma 1 we see from the condition (8) that we can construct the regular analytic function $P(x, y)(\partial \xi/\partial x + i\partial \xi/\partial y)$, where $P(x, y)$ is given by (11). Multiplying both sides of equation (12) and (13) by this analytic function and using the first condition of (7), we obtain

$$F'(z)P(x, y)G(z) = P(x, y)\left\{|G(z)|^2\frac{\partial}{\partial \xi} - iJ(z)\frac{\partial}{\partial \eta}\right\} \text{Re } F(z), \tag{14}$$

$$= P(x, y)\left\{J(z)\frac{\partial}{\partial \eta} + i|G(z)|^2\frac{\partial}{\partial \xi}\right\} \text{Im } F(z). \tag{15}$$

It can be seen from (14) and (15) that:

(i) If

$$\text{Re}\{F'(z)G(z)P(x, y)\} \neq 0, \qquad z \in D \cup \partial D, \tag{16}$$

then Re $F(z)$ is a monotonic function of ξ, and Im $F(z)$ is a monotonic function of η.

(ii) If

$$\text{Im}\{F'(z)G(z)P(x, y)\} \neq 0, \qquad z \in D \cup \partial D, \tag{17}$$

then Re $F(z)$ is a monotonic function of η, and Im $F(z)$ is a monotonic function of ξ. From Lemma 1 and the fact that $F'(z)$ is a regular analytic function, the expression in braces in (16) and (17) is also a regular analytic function when $z \in D \cup \partial D$. Since the real and imaginary parts of a regular analytic function are harmonic they cannot attain a maximum or minimum in the domain of analyticity, Titchmarsh [1, p. 167]. It follows therefore that if the expressions (16) and (17) do not vanish on the contour ∂D they cannot vanish in D. Hence the conditions (7) to (11) are sufficient to establish the monotonic properties of Re $F(z)$ and Im $F(z)$ given in Theorem 1.

COROLLARY TO THEOREM 1. *If $W(z)$ is a regular analytic function when $z \in D \cup \partial D$ and $W(z)$ does not vanish in this region, then* $\operatorname{Ln} W(z)$ *is also regular and analytic in $D \cup \partial D$. Letting $F(z) = \operatorname{Ln} W(z)$, we see that* $\operatorname{Re} F(z) = \operatorname{Ln} |W(z)|$, $\operatorname{Im} F(z) = \operatorname{Arg} W(z)$ *and*

$$\frac{\partial}{\partial \xi} \operatorname{Re} F(z) = \frac{1}{|W(z)|} \cdot \frac{\partial}{\partial \xi} |W(z)|, \qquad \frac{\partial}{\partial \eta} \operatorname{Re} F(z) = \frac{1}{|W(z)|} \cdot \frac{\partial}{\partial \eta} |W(z)|.$$

Therefore in $D \cup \partial D$, $\operatorname{Re} F(z)$ *will be a monotonic function of ξ or η if and only if $|W(z)|$ is a monotonic function of ξ or η. Hence in Theorem* 1 *if we assume the same transformation between x, y and ξ, η, and also the conditions* (7) *and* (8) *hold, and $W(z)$ is regular, analytic and nonzero in $D \cup \partial D$, then*:

(i) $|W(z)|$ *is a monotonic function of ξ, and* $\operatorname{Arg} W(z)$ *is a monotonic function of η in $D \cup \partial D$ if*

$$\operatorname{Re}\left\{\frac{W'(z)}{W(z)} G(z) P(x, y)\right\} \neq 0, \qquad z \in \partial D. \tag{9a}$$

(ii) $|W(z)|$ *is a monotonic function of η, and* $\operatorname{Arg} W(z)$ *is a monotonic function of ξ in $D \cup \partial D$ if*

$$\operatorname{Im}\left\{\frac{W'(z)}{W(z)} G(z) P(x, y)\right\} \neq 0, \qquad z \in \partial D, \tag{10a}$$

where in (9a) *and* (10a) *$P(x, y)$ is given by* (11).

3. Monotonic properties of $J_\nu(z)$. We now use Theorem 1 to prove the following theorem about $J_\nu(z)$.

THEOREM 2. *For complex z and $\nu \geqq 0$, $|J_\nu(z)|$ is a monotonic increasing function of $|y|$, and* $\operatorname{Arg} J_\nu(z)$ *is a monotonic increasing (decreasing) function of x for $y > 0$ $(y < 0)$.*

Proof. In the Corollary to Theorem 1 we let $W(z) = J_\nu(z)$, and choose ξ and η to correspond to the Cartesian coordinates $\xi = x$, $\eta = y$. The conditions (7) and (8) are satisfied and (11) gives $P(x, y) = 1$. $\{J_\nu(z)\}^{-1}$ and consequently $g_\nu(z)$ are regular analytic functions on and inside the contours $\partial D^{\pm}$ shown in Fig. 1. From the Corollary to Theorem 1 (compare (ii), (10a)) it can be seen that Theorem 2 is proved provided we can show that $\operatorname{Im}(g_\nu(z)) < 0$ on ∂D^+, and $\operatorname{Im}(g_\nu(z)) > 0$ on ∂D^-.

On the large semi-circles $\partial D_R^{\pm}(|z| \to \infty)$ the asymptotic expression for $J_\nu(z)$ as $|z| \to \infty$, see Watson [2, p. 199], gives

$$\operatorname{Im}(g_\nu(z)) \sim -\operatorname{sh} y \operatorname{ch} y \Big/ \left|\cos\left(z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right)\right|^2 \begin{cases} <0, & y \in \partial D_R^+, \\ >0, & y \in \partial D_R^-. \end{cases} \tag{18}$$

We now determine the sign of $\operatorname{Im}(g_\nu(z))$ on the straight lines $\partial D_\varepsilon^{\pm}$ given by $z = x \pm i\varepsilon$, $-\infty < x < \infty$, $\varepsilon > 0$. Putting $g_\nu(z)$ in the alternative form, Watson [2, p. 498],

$$g_\nu(z) = \frac{\nu}{z} - \frac{J_{\nu+1}(z)}{J_\nu(z)} = \frac{\nu}{z} + 2z \sum_{n=1}^{\infty} \frac{1}{(z^2 - j_{\nu n}^2)},$$

we find by direct differentiation that $g_\nu'(x) < 0$ for $-\infty < x < \infty$, $\nu \geqq 0$. Thus on the straight parts of the contours $\partial D_\varepsilon^{\pm}$ we have, for ε small,

$$g_\nu(x \pm i\varepsilon) = g_\nu(x) \pm i\varepsilon g_\nu'(x) + O(\varepsilon^2), \qquad \varepsilon < \delta(x),$$

where $\delta(x)$ is the radius of convergence of the power series, and equals the distance from x to the nearest pole of $g_\nu(z)$. This means that points x not near the real zeros of

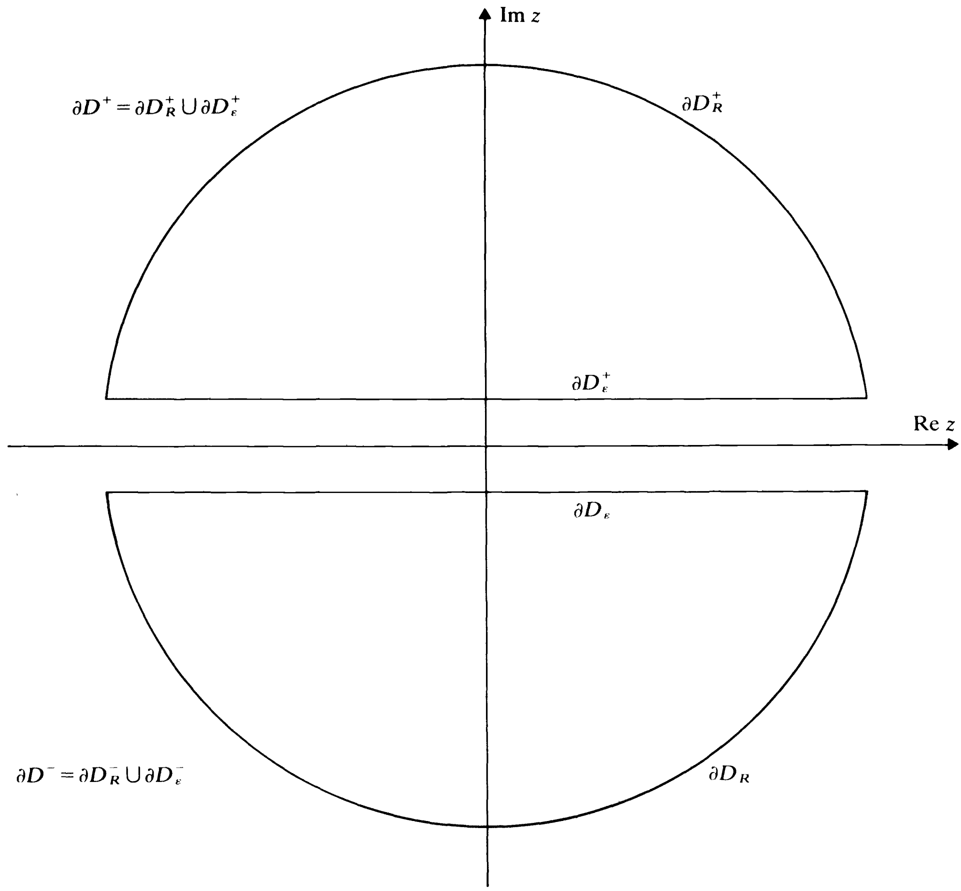

FIG. 1

$J_\nu(z)$ will satisfy the inequality $\delta(x) > \varepsilon$. To deal with the situation where x is near a zero of $J_\nu(z)$ and consequently $\delta(x) \leqq \varepsilon$, we indent the contour $\partial D_\varepsilon^\pm$ away from these zeros. Then near a zero $z = j_{\nu n}$, we have

$$g_\nu(z) = (z - j_{\nu n})^{-1}\{1 + (z - j_{\nu n}) J''_\nu(j_{\nu n})/(2J'_\nu(j_{\nu n})) + O((z - j_{\nu n})^2)\},$$

so that

$$\operatorname{Im} g_\nu(z) = -\rho^{-1} \sin \phi \cdot \{1 + O(\rho^2)\},$$

where $z - j_{\nu n} = \rho\, e^{i\phi}$.

Hence provided we let ε and ρ tend to zero together, with ρ always greater than $\sqrt{2}\varepsilon$ say, the equations above show that

$$\text{(19)} \qquad \operatorname{Im}(g_\nu(z)) \begin{cases} < 0, & y \in \partial D_\varepsilon^+, \\ > 0, & y \in \partial D_\varepsilon^-. \end{cases}$$

Hence inequalities (18) and (19) show that $\operatorname{Im}(g_\nu) < 0$ for $y > 0$ and $\operatorname{Im}(g_\nu(z)) > 0$ for $y < 0$. Therefore the proof of Theorem 2 is complete.

We note that for $\nu = n$, where n is an integer, the restriction that n be positive can be dropped. This follows from the relation $J_{-n}(z) = (-1)^n J_n(z)$.

COROLLARY.

$$|J_\nu(x \pm iy_1)| \leqq |J_\nu(x \pm iy_2)|, \qquad 0 \leqq y_1 \leqq y_2, \tag{20}$$

in particular

$$|J_\nu(x)| \leqq |J_\nu(z)|.$$

4. Monotonic properties of $H_\nu^{(1)}(z)$. In this section we prove two theorems which give monotonic properties of $H_\nu^{(1)}(z)$. Theorem 3 is similar to Theorem 2, with appropriate changes for $H_\nu^{(1)}(z)$. Theorem 4 is slightly different in that the cylindrical polar coordinates are used to establish monotonicity with respect to the radial and angular coordinates.

THEOREM 3. *$|H_\nu^{(1)}(z)|$ is a monotonic decreasing function of y, and* Arg $H_\nu^{(1)}(z)$ *is a monotonic increasing function of x, for $y \geqq 0$ and $\nu \geqq 0$.*

Proof. In the Corollary to Theorem 1 we let $W(z) = H_\nu^{(1)}(z)$, and choose $\xi = x$, $\eta = y$, so that the conditions (7) and (8) are satisfied and (11) gives $P(x, y) = 1$. Since $H_\nu^{(1)}(z)$ has no zeros or branch points in $y > 0$, $\nu \geqq 0$ the function $h_\nu(z)$ is regular and analytic in this domain. It follows from the corollary to Theorem 1 ((ii) expression (10a)) that Theorem 3 follows if and only if $\operatorname{Im}(h_\nu(z)) > 0$ on ∂D where ∂D is the contour shown in Fig. 2.

On the semi-circle $\partial D_R(|z| \to \infty,\ 0 < \operatorname{Arg} z < \pi)$, $h_\nu(z) \sim i$. On the small semi-circle $\partial D_\varepsilon(|z| = \varepsilon > 0, 0 < \operatorname{Arg} z < \pi)$, $h_\nu(z) \sim -\nu/z$ if $\nu > 0$ and $h_0(z) \sim (z \operatorname{Ln} z)^{-1}$. Finally on $\partial D_1(\varepsilon \leqq |z| < \infty, \operatorname{Arg} z = 0)$ and $\partial D_2(\varepsilon \leqq |z| < \infty, \operatorname{Arg} z = \pi)$, we have

$$\begin{aligned} \operatorname{Im}(h_\nu(z)) &= \frac{J_\nu(|x|)\,Y'_\nu(|x|) - J'_\nu(|x|)\,Y_\nu(|x|)}{J_\nu^2(|x|) + Y_\nu^2(|x|)} \\ &= 2\{\pi|x|\,|H_\nu^{(1)}(|x|)|^2\}^{-1}. \end{aligned} \tag{21}$$

The relationship $H_\nu^{(1)}(x\,e^{i\pi}) = -e^{i\nu\pi}H_\nu^{(2)}(x)$ is used in deriving (21) on ∂D_2. Thus as the contour expands to infinity we see that $\operatorname{Im}(h_\nu(z)) > 0$ for $\operatorname{Im}(z) \geqq 0$, $z \neq 0$. Hence Theorem 3 is proved.

COROLLARY.

$$|H_\nu^{(1)}(x + iy_2)| \leqq |H_\nu^{(1)}(x + iy_1)|, \qquad 0 \leqq y_1 \leqq y_2, \quad \nu \geqq 0,$$

$$|H_\nu^{(1)}(z)| \leqq |H_\nu^{(1)}(x)|, \qquad \operatorname{Im}(z) \geqq 0, \quad \nu \geqq 0.$$

THEOREM 4. *For complex $z(= r\,e^{i\theta})$ and $\nu \geqq 0$, $|H_\nu^{(1)}(z)|$ is a monotonic decreasing function of r, and* Arg $H_\nu^{(1)}(z)$ *is a monotonic decreasing function of θ, for $0 \leqq \theta \leqq \pi$, $0 < r < \infty$.*

Proof. In the Corollary to Theorem 1 we let $W(z) = H_\nu^{(1)}(z)$ and choose ξ and η to correspond to the cylindrical polar coordinates $\xi = r$, $\eta = \theta$, where $x = r\cos\theta$, $y = r\sin\theta$, $z = r\,e^{i\theta}$. It is not difficult to show that the conditions (7) and (8) are satisfied and that (11) gives $P(x, y) = (x^2 + y^2)^{1/2} = r$. Using these results and the Corollary to Theorem 1, (i) expression (9a), and also the fact that $h_\nu(z)$ is regular and analytic in $\operatorname{Im}(z) \geqq 0$, $z \neq 0$, we see that Theorem 4 is proved provided we can show that $\operatorname{Re}(zh_\nu(z)) < 0$ on ∂D. (See Fig. 2.)

On the semi-circle $\partial D_R(r \to \infty, 0 < \theta < \pi)$, we have $h_\nu(z) \sim i$ so that

$$\operatorname{Re}\{zh_\nu(z)\} \sim -r\sin\theta < 0, \qquad z \in \partial D_R. \tag{22}$$

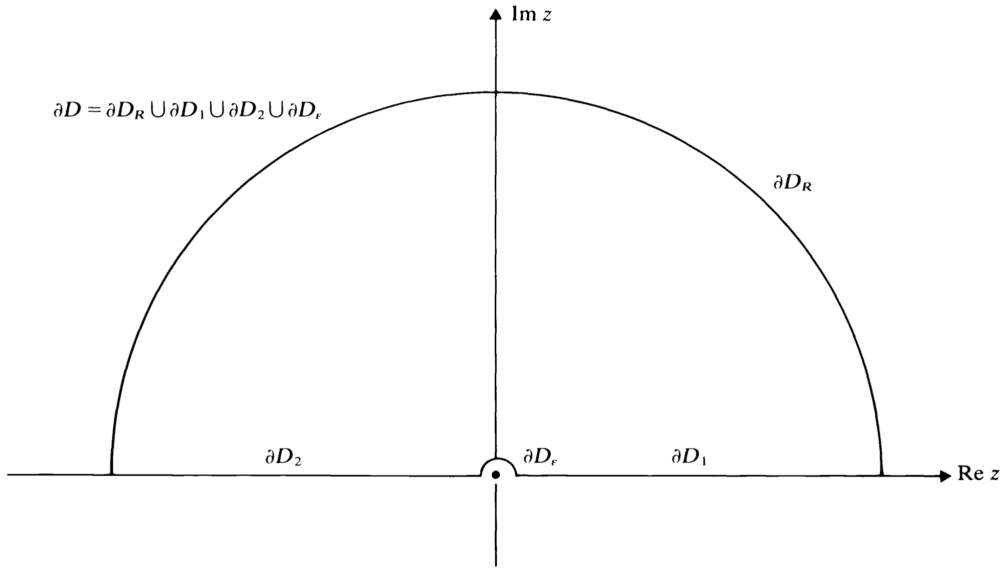

FIG. 2

On the small semi-circle, $\partial D_\varepsilon (r=\varepsilon>0, 0\leqq\theta\leqq\pi)$, we have $h_\nu(z)\sim-\nu/z$ if $\nu>0$ and $h_0(z)\sim(z\cdot\mathrm{Ln}\, z)^{-1}$, so that

$$\text{(23)}\qquad \begin{aligned} &\mathrm{Re}\,\{zh_\nu(z)\}\sim-\nu<0,\\ &\mathrm{Re}\,\{zh_0(z)\}\sim(\mathrm{Ln}\,\varepsilon)^{-1}<0, \qquad z\in\partial D_\varepsilon. \end{aligned}$$

On the lines $\partial D_1(\varepsilon\leqq r<\infty, \theta=0)$ and $\partial D_2(\varepsilon\leqq r<\infty, \theta=\pi)$ we have

$$\text{(24)}\qquad \begin{aligned} \mathrm{Re}\,\{zh_\nu(z)\}&=\frac{r(J'_\nu(r)J_\nu(r)+Y'_\nu(r)Y_\nu(r))}{J^2_\nu(r)+Y^2_\nu(r)}\\ &=\frac{r}{2}\cdot\frac{d}{dr}|H^{(1)}_\nu(r)|^2/|H^{(1)}_\nu(r)|^2, \qquad z\in\partial D_1\cup\partial D_2. \end{aligned}$$

Now from Watson [2, p. 444] we have

$$\text{(25)}\qquad |H^{(1)}_\nu(r)|^2=\frac{8}{\pi^2}\int_0^\infty K_0(2r\,\mathrm{sh}\,t)\,\mathrm{ch}\,2\nu t\,dt>0, \qquad 0<r<\infty,$$

and

$$\text{(26)}\qquad \frac{d}{dr}|H^{(1)}_\nu(r)|^2=-\frac{8}{\pi^2}\int_0^\infty 2\,\mathrm{sh}\,t\,\mathrm{ch}\,2\nu t K_1(2r\,\mathrm{sh}\,t)\,dt<0, \qquad 0<r<\infty.$$

From (25) and (26) we see that

$$\text{(27)}\qquad \mathrm{Re}\,(zh_\nu(z))<0, \qquad z\in\partial D_1\cup\partial D_2.$$

Thus the relations (22), (23) and (27) show that $\mathrm{Re}\,(zh_\nu(z))<0$ on ∂D. Theorem 4 is therefore proved on letting the contour ∂D expand to infinity.

Acknowledgment. We would like to thank the referees for helpful comments on the presentation of this paper.

REFERENCES

[1] E. C. TITCHMARSH, *The Theory of Functions*, 2nd ed., Oxford University Press, Oxford, 1939.
[2] G. N. WATSON, *Theory of Bessel functions*, 2nd ed., Cambridge University Press, Cambridge, 1944.

SIAM. J. MATH. ANAL.
Vol. 9, No. 6, December 1978

© 1978 Society for Industrial and Applied Mathematics
0036-1410/78/0906-0016 $01.00/0

A FREE BOUNDARY OPTIMIZATION PROBLEM*

ANDREW ACKER†

Abstract. Given a convex set $Q \subset R^2$ (bounded by a simple closed curve) and a constant $A > 0$, we determine the doubly-connected region Ω encircling (but not intersecting) Q, with area $|\Omega| \leqq A$, which has the least capacitance.

1. Notation. Ω is a doubly-connected region in R^2 with simple, closed curves Γ^* and Γ as (respectively) inner and outer boundary components. (See Fig. 1.) $S^* \supset \Gamma^*$ and $S \supset \Gamma$ are the finite and infinite components of $R^2 \backslash \Omega$. The capacitance K of Ω is defined by $K = \int_\gamma |\nabla U(p)| \cdot |dp|$, where the electrostatic potential $U(p)$ is the unique continuous function on R^2 which is harmonic in Ω and satisfies: $U = 1$ in S^*, $U = 0$ in S, and where $\gamma \subset \Omega$ is an equipotential curve of U. The notation Γ_c^*, $U_c(p)$, $\bar{S}$, $\dot{S}_i^*$, $\dot{K}_i, \cdots$ refers to other regions Ω_c, $\bar{\Omega}$, $\dot{\Omega}_i, \cdots$ with (unless otherwise stated) the same properties as Ω.

2. Introduction and main results. We will investigate the following free boundary optimization problem.

Problem 1. (See Fig. 1.) *Given a compact set $Q \subset R^2$ (whose boundary ∂Q is a simple, closed curve) and a constant $A > 0$, we seek the region Ω which minimizes K subject to the constraints that $S^* \supset Q$ and $|\Omega| \leqq A$ (where $|\Omega|$ refers to area).*

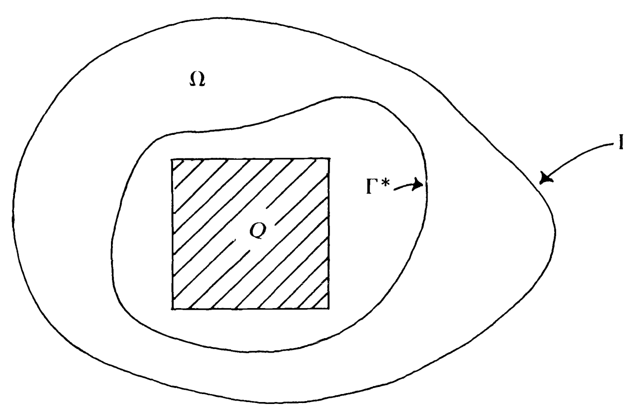

FIG. 1. *What doubly-connected region Ω of area A encircling Q has the least capacitance?*

If $Q = \{p \in R^2 : |p| \leqq r_0\}$, $r_0 > 0$, then the solution of Problem 1 at area $A = \pi(r_1^2 - r_0^2) > 0$ is the annulus $\Omega = \{p \in R^2 : r_0 < |p| < r_1\}$, as follows from the well known isoperimetric inequality of T. Carleman [10].

We will solve Problem 1 when Q is convex. In this case, the solutions over all $A > 0$ are given by the monotone family of regions $\{\Omega_c : c > 0\}$ defined in the following theorem due to D. E. Tepper [18], [19].

* Received by the editors June 29, 1976, and in final revised form January 30, 1978. Presented at the conference on partial differential equations, Oberwolfach, February, 1977. An abstract appeared in Notices of the American Mathematical Society, 23 (1976), p. A-645.

† Mathematisches Institut I, Universität Karlsruhe (TH), 75 Karlsruhe 1, Englerstrasse 2, Postfach 6380, Federal Republic of Germany.

THEOREM 1. *If Q is starlike relative to $p_0 \in Q$, then*:

(a) *For any $c > 0$, there exists a region Ω_c such that $S_c^* = Q$ and $|\nabla U_c(p)| = c$ on Γ_c.*

(b) *$\Omega_{c'} \subset \Omega_c$ whenever $c' \geqq c > 0$.*

(c) *$\bigcup_{c>0} \Gamma_c = R^2 \setminus Q$.*

(d) *$Q \cup \Omega_c$ is starlike relative to p_0* for each $c > 0$. *Moreover, if Q is convex, then $Q \cup \Omega_c$ is convex for each $c > 0$.*

Here, for any $p \in \Gamma_c$ we define $|\nabla U_c(p)| := \lim_{q \to p} |\nabla U_c(q)|$ $(q \in \Omega_c)$ if the limit exists. $|\nabla U_c(p)|$ represents the surface charge density induced by the potential $U_c(p)$ at $p \in \Gamma_c$.

Our main result is the following theorem:

THEOREM 2. *Let Q be starlike, and for each $A > 0$, let $c > 0$ be the unique value such that $|\Omega_c| = A$. We consider two cases*:

Case 1: *If Q is convex, then*

$$K > K_c \tag{1}$$

for any region $\Omega \neq \Omega_c$ such that $S^* \supset Q$ and $|\Omega| \leqq A$. Thus Ω_c uniquely solves Problem 1.

Case 2. *If Q is not convex, but ∂Q has bounded curvature, then Ω_c does not solve Problem* 1 *if $A > 0$ is sufficiently small.*

The following is an equivalent alternative formulation of Theorem 2 in the context of analytic function theory. (We set $z = x + iy = (x, y) = p$.)

THEOREM 3. *Let Q be starlike, and for any $c > 0$ let $\{\Omega\}_c$ be the class of all regions Ω which are conformally equivalent to Ω_c and satisfy $S^* \supset Q$. If Q is convex, then Ω_c is uniquely area minimizing in $\{\Omega\}_c$ for any $c > 0$, i.e.*

$$|\Omega| > |\Omega_c| \tag{2}$$

for any region $\Omega \neq \Omega_c$ belonging to $\{\Omega\}_c$. If Q is not convex, but ∂Q has bounded curvature, then Ω_c is not area minimizing in $\{\Omega\}_c$ for $c > 0$ sufficiently large.

Remark 1. Since K_c and $|\Omega_c|$ can (in principle) be determined for any $c > 0$, (1) and (2) are isoperimetric inequalities.

Remark 2. A variation of Problem 1 consists of minimizing the capacitance in the class $\{\Omega\}_A$ of regions Ω satisfying $S^* = Q$ and $|\Omega| \leqq A$ (i.e. separation of Ω away from the geometric constraint Q is prohibited). In [1], the author showed that the region Ω_c satisfying $|\Omega_c| = A$ is uniquely capacitance minimizing in $\{\Omega\}_A$ whenever Q is starlike.

Remark 3. For any Q, a necessary condition for a sufficiently regular region Ω to solve Problem 1 can be obtained using the following variational formula due to H. Poincaré. (See [16].) For a small, smooth perturbation of a sufficiently regular region Ω, the variation in capacitance is given to first order by

$$\delta K \simeq \int_{\Gamma^* \cup \Gamma} |\nabla U(p)|^2 \delta n(p) \cdot |dp|, \tag{3}$$

where $\delta n(p)$ is the shift in the boundary of Ω in the direction of the interior normal at $p \in \Gamma^* \cup \Gamma$. Since the variation in $|\Omega|$ is given to first order by $\delta|\Omega| \simeq -\int_{\Gamma^* \cup \Gamma} \delta n(p) \cdot |dp|$, it follows that Ω (sufficiently regular) can solve Problem 1 only if $|\Omega| = A$ and

$$|\nabla U(p)| = c \quad \text{on } (\Gamma^* \cup \Gamma) \setminus Q \quad \text{and} \quad |\nabla U(p)| \geqq c \quad \text{on } \Gamma^* \cap \partial Q \tag{4}$$

for some $c > 0$. If Q is convex and ∂Q has bounded curvature, then the region Ω_c solves (4) for each $c > 0$. Apparently the free boundary problem (4) has not been investigated in the general case.

The proof of Theorem 2, Case 2 is in the following section. The proof in § 5 of case 1 is based on a particular, nearly continuous, area preserving deformation of Ω_c into any region Ω satisfying $|\Omega| = |\Omega_c|$. We show, essentially by applying (3), that capacitance is nondecreasing throughout the deformation.

3. Further properties of the free boundary solutions. The following lemma is closely related to Lindelöf's principle and Montel's principle. (See [11, Chap. 1].)

LEMMA 4. *Let Ω and $\bar{\Omega}$ be regions such that $S^* \subset \bar{S}^*$ and $S \supset \bar{S}$. Then*:

(a) *If $p \in \Gamma \cap \bar{\Gamma}$, then $|\nabla U(p)| \leqq |\nabla \bar{U}(p)|$ if both derivatives exist.*

(b) *If $p \in \Gamma^* \cap \bar{\Gamma}^*$, then $|\nabla U(p)| \geqq |\nabla \bar{U}(p)|$ if both derivatives exist.*

Proof. The maximum principle for harmonic functions implies that $0 \leqq \tilde{U}(q) \leqq 1$, $q \in R^2$, where $\tilde{\Omega} := R^2 \backslash (S^* \cup \bar{S})$. Further application of the maximum principle leads to the inequalities: $U(q) \leqq \tilde{U}(q) \leqq \bar{U}(q)$, $q \in R^2$. If $W(q) = \bar{U}(q) - U(q)$, then $W(p) = 0$ if $p \in \Gamma \cap \bar{\Gamma}$ or $p \in \Gamma^* \cap \bar{\Gamma}^*$. The results follow from this.

LEMMA 5. *Let Q be starlike. If $\{\Omega_c: c > 0\}$ are the regions defined in Theorem* 1, *then*:

(a) *For each $c > 0$, Γ_c is an analytic curve.*

(b) *If Q is convex, then $Q \cup \Omega_c(\lambda, 1)$ is convex for all $c > 0$ and $0 < \lambda < 1$, where $\Omega_c(\lambda, 1) = \{p \in \Omega_c: \lambda < U_c(p) < 1\}$.*[1]

(c) *If Q is convex, then $|\nabla U_c(p)|$ is strictly monotone increasing with increasing $U_c(p)$ along all curves of steepest ascent (of $U_c(p)$). Thus $|\nabla U_c(p)| > c$ in Ω_c.*

(d) *If Q is convex, and the curvature of ∂Q is bounded by a constant $B < \infty$, then the curvature of Γ_c has the same bound B, $c > 0$.*

(e) *Assume ∂Q has bounded curvature, and let θ be the arc length of ∂Q. If the values $A > 0$ and $c > 0$ are related by $A = |\Omega_c|$, then there exist constants $A_0 > 0$, $c_0 < \infty$, and $M < \infty$ such that the following inequalities* ((5)–(10)) *hold whenever $0 < A \leqq A_0$, or $c \geqq c_0$.*

$$|\underset{\sim}{d}(p, Q) - (1/c)| \leqq (M/c^2) \tag{5}$$

for all $p \in \Gamma_c$, where $\underset{\sim}{d}(p, Q) = \inf \{|p - q)|: q \in Q\}$.

$$||\nabla U_c(p)| - c| \leqq M \tag{6}$$

for all $p \in \partial Q$.

$$|A - (\theta/c)| \leqq (M/c^2), \tag{7}$$

$$|c - (\theta/A)| \leqq M, \tag{8}$$

$$|K_c - c \cdot \theta| \leqq M, \tag{9}$$

$$|K_c - (\theta^2/A)| \leqq M. \tag{10}$$

Proof. (*Part* (a).) For c fixed, one can define the harmonic conjugate $V(p)$ of $U_c(p)$ on $\Omega_c \backslash \gamma$ (where γ is a Jordan slit connecting ∂Q and Γ_c) in such a way that $F(z) = U_c + iV$ maps $\Omega_c \backslash \gamma$ conformally onto $(0, 1) \times (0, K_c)$. If $G = F^{-1}$, then $G(z)$ has a K_c-periodic analytic extension to $(0, 1) \times R$. The condition that $|\nabla U_c(p)| = c$ on Γ_c implies that a continuous extension of $|G'(z)|$ to $[0, 1) \times R$ exists and satisfies $|G'(z)| = (1/c)$ on $\{0\} \times R$. Therefore $\log (c \cdot |G'(z)|)$ has an antisymmetric harmonic continuation to $(-1, 1) \times R$. Thus $G(z)$ can be analytically continued to $(-1, 1) \times R$, and Γ_c (the image of $\{0\} \times [0, 1]$ under G) is an analytic curve.

[1] In general, we define $\Omega(\alpha, \beta) = \{p \in \Omega: \alpha < U(p) < \beta\}$ for any $0 \leqq \alpha < \beta \leqq 1$.

Parts (b) *and* (c). If the word "strong" is dropped in (c), then both results follow using Tepper's proof (in [18]) that $Q \cup \Omega_c$ is convex. The general case of (c) follows by applying the strong maximum principle to $D_x \log(|G'(x+iy)|)$ on $(\lambda_1, \lambda_2)\times R$, $0<\lambda_1<\lambda_2<1$.

Part (d). Define $W(x, y)=D_y \arg(G'(x+iy))$. The bound on the curvature of ∂Q is equivalent to the condition: $|W(1, y)| \leqq B\cdot|G'(1+iy)|$ for all $y \in R$. Part (c) implies that $|G'(1+iy)| \leqq |G'(0+iy)|=(1/c)$, so that $|W(1, y)| \leqq (B/c)$ for all y. Since $W(x, y)$ is harmonic and $D_x W(0, y)=0$, $y \in R$, the maximum principle implies that $|W(0, y)| \leqq (B/c)=B\cdot|G'(0+iy)|$ for all y, from which the bound on the curvature of Γ_c follows.

Part (e). Let $r=(1/B)$ be the minimum radius of curvature of ∂Q. For a fixed point $p_0 \in \partial Q$, let $\bar{p} \in Q$ and $\tilde{p} \in R^2\backslash Q$ be the centers of the circles of radius r tangent to ∂Q at p_0. If $\bar{\Omega}_c=\{p \in R^2: r<|p-\bar{p}|<\bar{r}_c\}$ is defined such that $|\nabla \bar{U}_c(p)|=c$ on $\bar{\Gamma}_c$, then elementary calculations show that $c \leqq |\nabla \bar{U}_c(p)| \leqq c+B$ in $\bar{\Omega}_c$ and $(c+B)^{-1} \leqq \bar{r}_c-r \leqq c^{-1}$ for each $c>0$. For each $c>e\cdot B$ (where $e=\exp(1)$), a unique region $\tilde{\Omega}_c=\{p \in R^2: \tilde{r}_c<|p-\tilde{p}|<r\}$ exists such that $(r/e)<\tilde{r}_c<r$ and $|\nabla \tilde{U}_c(p)|=c$ on $\tilde{\Gamma}_c^*$. Again, elementary calculations show that $c^{-1}<r-\tilde{r}_c<(c-e\cdot B)^{-1}$ and that $c\cdot(1-B\cdot(c-e\cdot B)^{-1}) \leqq |\nabla \tilde{U}_c(p)| \leqq c$ in $\tilde{\Omega}_c$. Since $\Gamma_c \to \partial Q$ uniformly as $c \to +\infty$, there is a constant $c_1<\infty$ such that if $c>c_1$, then $\underline{d}(p, Q)<(1-e^{-1})\cdot r$ for all $p \in \Gamma_c$. For $c>c_1$, Lemma 4 shows that $\tilde{S}_c^* \subset S_c \subset \bar{S}_c$, from which it follows that $(c+B)^{-1} \leqq \underline{d}(p, Q) \leqq (c-e\cdot B)^{-1}$, where p is the intersection of Γ_c with the perpendicular to ∂Q through p_0. Since $p_0 \in \partial Q$ is arbitrary, we obtain $|\underline{d}(p, Q)-c^{-1}| \leqq 2(1+e)\cdot(B/c^2)$ for $c>c_0=\max\{c_1, 2e\cdot B\}$ for all $p \in \Gamma_c$. Now Lemma 4 implies that $c\cdot(1-B\cdot(c-e\cdot B)^{-1}) \leqq |\nabla \tilde{U}_c(p_0)| \leqq |\nabla U_c(p_0)| \leqq |\nabla \bar{U}_c(p_0)| \leqq c+B$, from which it follows that $||\nabla U_c(p)|-c| \leqq 3B$ for all $p \in \partial Q$ and $c>c_0$. This completes the proof of (5) and (6). (7) is easily derived from (5), and (8) follows from (7). Further, (6) implies (9) and (8) and (9) imply (10).

Proof of Theorem 2, *Case* 2. Let $\hat{Q} \neq Q$ be the convex hull of Q, and $\hat{\theta}$ be the length of the boundary of $\hat{Q}$. We will show that $\hat{K}<K_c$ for $A>0$ sufficiently small, where $\hat{\Omega}$ is defined such that $\hat{S}^*=\hat{Q}$, $|\hat{\Omega}|=A$, and $|\nabla \hat{U}(p)|$ is a constant on $\hat{\Gamma}$. Applying (10) to $\hat{\Omega}$, we obtain

$$|\hat{K}-(\tilde{\theta}^2/A)| \leqq M \quad (A>0 \text{ small}). \tag{11}$$

Since $0<\hat{\theta}<\theta$, it follows from (10) and (11) that $\hat{K}<K_c$ when $A>0$ is sufficiently small.

Remark 4. In [1], the region Ω_c was interpreted as an optimally shaped insulation layer about a starlike interior Q. In this context, (5) shows that if ∂Q has bounded curvature, then the optimal insulation layer of sufficiently small area is essentially uniformly thick.

4. A special monotone sequence of regions. The purpose of this section is to define a special sequence of regions and to derive properties of these regions sufficient to form the basis for the proof of Theorem 2, Case 2 in the following section. The following notation involving distances will be used: $\underline{d}(p, \Gamma)=\inf\{|p-q|: q \in \Gamma\}$, $\bar{d}(p, \Gamma)=\sup\{|p-q|: q \in \Gamma\}$, $\underline{d}(\Gamma^*, \Gamma)=\inf\{|p-q|: p \in \Gamma^*, q \in \Gamma\}$, and $\bar{d}(\Gamma^*, \Gamma)=\sup\{\underline{d}(p, \Gamma^*): p \in \Gamma\}$. Notice that $\bar{d}(\Gamma^*, \Gamma) \neq \bar{d}(\Gamma, \Gamma^*)$.

LEMMA 6. *Let Q be convex, and let $\Omega_{\hat{c}}$ and Ω be regions such that $S_{\hat{c}}^*=Q$, $|\nabla U_{\hat{c}}(p)|=\hat{c}$ on $\Gamma_{\hat{c}}$, $S^* \supset Q$, and $A_0 := |S_{\hat{c}}\backslash S|=|S^*\backslash Q|$. Then for any $\delta>0$ and $n \in N$ satisfying $n\cdot\delta=A_0$, there exists a sequence of regions $\dot{\Omega}_i$, $i=0, \cdots, n$, with the following properties*:

(a) $\dot{\Omega}_0=\Omega_{\hat{c}}$.

(b) $\dot{S}_i^* \subset \dot{S}_{i+1}^*$ *and* $\dot{S}_i \supset \dot{S}_{i+1}$, $i = 0, \cdots, n-1$.

(c) $\dot{S}_i^*$ *and* $\dot{S}_i^* \cup \dot{\Omega}_i$ *are convex sets for each i.*

(d) *For each i, there is a constant* $c_i > 0$ *such that* $|\nabla \dot{U}_i(p)| = c_i$ *on* $\dot{\Gamma}_i$.

(e) $|(\dot{S}_{i+1}^* \backslash \dot{S}_i^*) \cap S^*| = |(\dot{S}_i \backslash \dot{S}_{i+1}) \backslash S| = \delta$, $i = 0, \cdots, n-1$. *Thus* $|(\dot{S}_i^* \backslash Q) \cap S^*| = |(S_{\hat{c}} \backslash \dot{S}_i) \backslash S| = i \cdot \delta$ *for each i.*

Proof. (See Fig. 2.) We define the $\dot{\Omega}_i$ inductively. $\dot{\Omega}_0$ is defined by (a). Assume now for $0 \leqq m < n = (A_0/\delta)$ that $\dot{\Omega}_i$, $i = 0, \cdots, m$, have been defined and satisfy (a)–(e). Since $|\dot{S}_m \backslash S| = |S_{\hat{c}} \backslash S| - |(S_{\hat{c}} \backslash \dot{S}_m) \backslash S| = A_0 - m \cdot \delta \geqq \delta$, there is by Theorem 1 a unique smallest region $\mathring{\Omega}_m \supset \dot{\Omega}_m$ such that $\mathring{S}_m^* = \dot{S}_m^*$, $|\nabla \mathring{U}_m(p)|$ is constant on $\mathring{\Gamma}_m$, and $|(\dot{S}_m \backslash \mathring{S}_m) \backslash S| = \delta$. Moreover $\mathring{S}_m^* \cup \mathring{\Omega}_m$ is convex. Now $S^* \cap \mathring{S}_m \subset \mathring{S}_m \backslash S$, and in fact $|S^* \cap \mathring{S}_m| < |\mathring{S}_m \backslash S| = A_0 - (m+1) \cdot \delta$. Therefore $|\mathring{\Omega}_m \cap S^*| = |S^* \backslash Q| - |\mathring{S}_m \cap S^*| - |(S_m^* \backslash Q) \cap bS^*| > \delta$. Thus there is a unique largest $0 < \lambda < 1$ such that $|\mathring{\Omega}_m(\lambda, 1) \cap S^*| = \delta$. We define $\dot{\Omega}_{m+1} = \mathring{\Omega}_m(0, \lambda)$. Then $\dot{S}_{m+1}^*$ is convex by Lemma 5(b) and $|\nabla \dot{U}_{m+1}(p)| = (1/\lambda) \cdot |\nabla \mathring{U}_m(p)|$ is constant on $\dot{\Gamma}_{m+1} = \mathring{\Gamma}_m$. Thus, the appropriate $(m+1)$st region has been defined. (Remark: The words "smallest" and "largest" can be omitted except when $m = n-1$.)

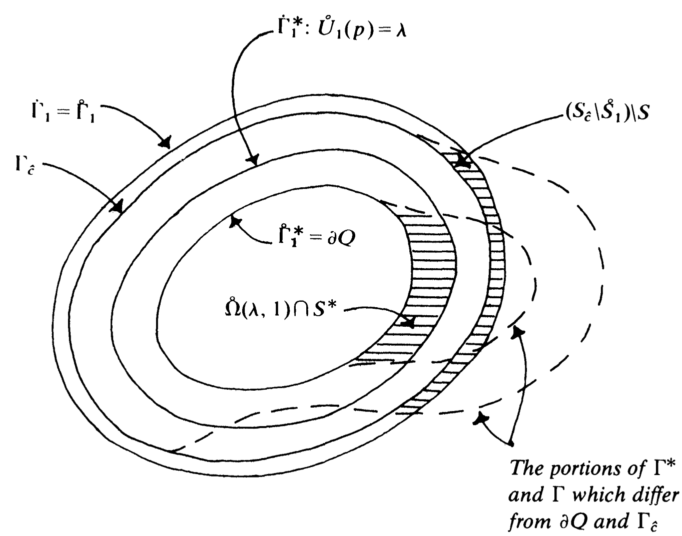

FIG. 2. *The definition of* $\dot{\Omega}_1$. *First define* $\mathring{\Omega}_1 \supset \Omega_{\hat{c}}$ *such that* $\mathring{S}_1^* = Q$, $|(S_{\hat{c}} \backslash \mathring{S}_1) \backslash S| = \delta$, *and* $|\nabla \mathring{U}_1(p)|$ *is constant on* $\mathring{\Gamma}_1$. *Then define* $\dot{\Omega}_1 = \mathring{\Omega}_1(0, \lambda)$, *where* $0 < \lambda < 1$ *is chosen such that* $|S^* \cap \mathring{\Omega}_1(\lambda, 1)| = \delta$. *One defines* $\dot{\Omega}_{i+1}$ *from* $\dot{\Omega}_i$, $i = 1, \cdots, n-1$, *in the same way.*

LEMMA 7. *Let* $\Omega_{\hat{c}}$ *and* Ω *have the properties assumed in Lemma 6. If* ∂Q *has bounded curvature, then the regions* $\dot{\Omega}_i$, $i = 0, \cdots, n$, *defined in Lemma 6 have the following properties.*

(a) *There exist constants* $\underline{C} > 0$, $\bar{d} < \infty$, *and* $\underline{d} > 0$ *such that for any* $\delta > 0$ *and* $n \in N$ *satisfying* $n \cdot \delta = A_0$, *we have*:

$$|\nabla \dot{U}_i(p)| \geqq \underline{C} \text{ on } \dot{\Omega}_i, \tag{12}$$

$$\bar{d}(\dot{\Gamma}_i^*, \dot{\Gamma}_i) \leqq \bar{d}, \quad \text{and} \tag{13}$$

$$\underline{d}(\dot{\Gamma}_i^*, \dot{\Gamma}_i) \geqq \underline{d}, \tag{14}$$

for $i = 0, \cdots, n$.

(b) *For any* $\alpha \in [0, A_0)$, *there exist constants* $\bar{C}(\alpha) < \infty$, $M(\alpha) < \infty$, *and* $\delta(\alpha) > 0$ *such that for any* $\delta \in (0, \delta(\alpha)]$ *we have*:

$$|\nabla \dot{U}_i(p)| \leqq \bar{C}(\alpha) \quad \text{in } \dot{\Omega}_i \quad \text{if} \quad i \cdot \delta \leqq \alpha, \tag{15}$$

$$c_{i+1} - c_i \leqq M(\alpha) \cdot \delta \quad \text{if} \quad (i+1) \cdot \delta \leqq \alpha, \tag{16}$$

$$\bar{d}(\dot{\Gamma}_i^*, \dot{\Gamma}_{i+1}^*) \leqq M(\alpha) \cdot \delta \quad \text{if} \quad (i+1) \cdot \delta \leqq \alpha, \tag{17}$$

$$\bar{d}(\dot{\Gamma}_i, \dot{\Gamma}_{i+1}) \leqq M(\alpha) \cdot \delta \quad \text{if} \quad (i+1) \cdot \delta \leqq \alpha, \quad \text{and} \tag{18}$$

$$\big| |\nabla \dot{U}_i(p)| - |\nabla \dot{U}_i(q)| \big| \leqq M(\alpha) \cdot |p-q| \quad \text{for all } p, q \in \dot{\Omega}_i(0, \tfrac{1}{2}) \quad \text{if} \quad i \cdot \delta \leqq \alpha. \tag{19}$$

Proof. For the proof of Part (a), let ω be a region such that $s \subset S_{\hat{c}}$, $s^* \supset Q$, s^* is convex, $|\nabla u(p)| = c$ on γ, $|(s^* \backslash Q) \cap S^*| = |(S_{\hat{c}} \backslash s) \backslash S|$, $\gamma^* \cap S^* \neq \varnothing$, and $\gamma \cap (\Omega \cup \Gamma) \neq \varnothing$. (All conditions are satisfied when $\omega = \dot{\Omega}_i$ for some i.) We assume throughout that 0 (the origin) is located in the interior of Q. To prove (12), let $0 < \underline{r} = \underline{d}(0, \partial Q) < r = \underline{d}(0, \gamma) \leqq \bar{r} = \bar{d}(0, \Gamma)$. If $\tilde{\Omega} = \{p \in R^2 : \underline{r} < |p| < r\}$, then $|\nabla \tilde{U}(p)| = (r \cdot \log(r/\underline{r}))^{-1} \leqq (\bar{r} \cdot \log(\bar{r}/\underline{r}))^{-1}$ for all $p \in \tilde{\Gamma}$. Further, if $p \in \tilde{\Gamma} \cap \gamma$, then $c = |\nabla u(p)| \geqq |\nabla \tilde{U}(p)|$ by Lemma 4. Thus, in (12) it suffices to set $\underline{C} = (\bar{r} \cdot \log(\bar{r}/\underline{r}))^{-1}$.

For the proof of (13), let l be the curve of steepest ascent of $u(q)$ beginning at $p \in \gamma$. Since (by Lemma 5(c)) $|\nabla u(q)| > c$ in ω, we have $c \cdot L(l) < \int_l |\nabla u(q)| \cdot |dq| = 1$, where $L(\cdot)$ denotes arc length. Thus $\underline{d}(p, \gamma^*) \leqq L(l) \leqq (1/c)$. Since $p \in \gamma$ is arbitrary, we have $\bar{d}(\gamma^*, \gamma) \leqq (1/c) \leqq (1/\underline{C})$, proving (13).

For the proof of (14), let $\varepsilon = \underline{d}(\gamma^*, \gamma) > 0$. There exist points $p \in \gamma$ and $q \in \gamma^*$ such that $|p-q| = \varepsilon$. If $\gamma_0^* = \{\lambda \cdot q : 1 - (\underline{r}/|q|) \leqq \lambda \leqq 1\}$ and $\omega_\varepsilon = \{p \in R^2 : 0 < \underline{d}(p, \gamma_0^*) < \varepsilon\}$, then $p \in \gamma \cap \gamma_\varepsilon$ and $c = |\nabla u(p)| \geqq |\nabla u_\varepsilon(p)|$ by Lemma 4. On the other hand, $\inf\{|\nabla u_\varepsilon(p)| : p \in \gamma_\varepsilon\} \to +\infty$ as $\varepsilon \to 0+$. Thus, a uniform lower bound for $\underline{d}(\gamma^*, \gamma)$ follows if a uniform upper bound for c is established. Since $c < (1/\bar{d}(\gamma^*, \gamma))$, it suffices to find a positive lower bound for $\bar{d}(\gamma^*, \gamma)$. To this end, we note that there exists a constant $\eta > 0$ such that the area of the portion of Ω on either side of any straight line intersecting S^* is at least η. The convexity of s^* implies that $|(\omega \cup s) \backslash S| \geqq |S^* \cap (\omega \cup s)| + \eta$. Since $|s \backslash S| = A_0 - |(S_{\hat{c}} \backslash s) \backslash S| = A_0 - |(s^* \backslash Q) \cap S^*| = |S^* \cap (\omega \cup s)|$, we have $|\omega \backslash S| \geqq \eta$. If μ is the length of the boundary of the convex hull of $S^* \cup \Omega$, then $L(\gamma^* \backslash S) \leqq \mu$ and $L(\gamma \backslash S) \leqq \mu$. In fact $|\omega \backslash S| \leqq \mu \cdot \bar{d}(\gamma^*, \gamma)$. It follows that $\bar{d}(\gamma^*, \gamma) \leqq (\eta/\mu)$, proving (14).

Part (b). For the proof of (15), let $\bar{C}_i = \sup\{|\nabla \dot{U}_i(p)| : p \in \dot{\Gamma}_i^*\}$, $i = 0, \cdots, n$. $\bar{C}_0$ is finite due to the bounded curvature of $\dot{\Gamma}_0^* = \partial Q$. For fixed i, $V(p) := |\nabla \mathring{U}_i(p)|$ is a subharmonic function in $\mathring{\Omega}_i$ whose continuous extension to the boundary satisfies $V(p) \leqq \bar{C}_i$ on Γ_i^* and $V(p) \leqq c_i \leqq \bar{C}$ on $\mathring{\Gamma}_i$, where $\bar{C}$ is the bound for the constants c_i established in the proof of (14). Therefore

$$|\nabla \mathring{U}_i(p)| \leqq (\bar{C}_i - \bar{C}) \cdot \mathring{U}_i(p) + \bar{C}$$

in $\mathring{\Omega}_i$. Now $\dot{\Gamma}_{i+1}^* = \{p \in R^2 : \mathring{U}_i(p) = 1 - \varepsilon\}$, where $0 < \varepsilon < 1$ is appropriately chosen. Thus, the above inequality implies that

$$\bar{C}_{i+1} \leqq \bar{C}_i + \bar{C} \cdot ((1-\varepsilon)^{-1} - 1).$$

Now choose $p \in \dot{\Gamma}_i^*$ and $q \in \dot{\Gamma}_{i+1}^*$ such that $|p - q| = \underline{d}(\dot{\Gamma}_i^*, \dot{\Gamma}_{i+1}^*)$. Then $\varepsilon \leqq \int_l |\nabla \mathring{U}_i(p')| \cdot |dp'| \leqq \bar{C}_i \cdot \underline{d}(\dot{\Gamma}_i^*, \dot{\Gamma}_{i+1}^*)$, where l is the straight line joining p and q. Further, for any $\alpha \in [0, A_0)$, there is a constant $N(\alpha) > 0$ with the property that for any region $\omega \subset S^* \backslash Q$ of area $|\omega| \geqq A_0 - \alpha$, whose boundary is rectifiable, the length of the boundary of ω relative to S^* exceeds $N(\alpha)$. If $(i+1)\cdot\delta \leqq \alpha$, then $N(\alpha)\cdot \underline{d}(\dot{\Gamma}_i^*, \dot{\Gamma}_{i+1}^*) \leqq \delta$. Thus

$$\varepsilon \leqq (\bar{C}_i/N(\alpha))\cdot\delta, \tag{20}$$

and

$$\bar{C}_{i+1} - \bar{C}_i \leqq \bar{C}\cdot((1-(\delta/N(\alpha))\cdot\bar{C}_i)^{-1}-1)$$

for $(i+1)\cdot\delta\leqq\alpha$. Since this difference inequality corresponds to the differential inequality

$$D_x\bar{C}(x) \leqq (\bar{C}/N(\alpha))\cdot\bar{C}(x), \qquad 0\leqq x\leqq\alpha,$$

it is clear that the constants $\bar{C}_i$, $i\cdot\delta\leqq\alpha$, are uniformly bounded over all sufficiently small $\delta>0$.

For the proof of (16), we have $c_{i+1}\leqq(c_i/(1-\varepsilon))$, $i=0,\cdots,n-1$, where $\mathring{U}_i(p)=1-\varepsilon$ on $\dot{\Gamma}_{i+1}^*$. Thus for $\delta>0$ sufficiently small and for $i\cdot\delta\leqq\alpha$, it follows using (15) and (20) that

$$c_{i+1}-c_i \leqq 2c_i\cdot\varepsilon \leqq 2\bar{C}(\alpha)\cdot\varepsilon \leqq 2((\bar{C}(\alpha))^2/N(\alpha))\cdot\delta.$$

To prove (17), we note that (as in the proof of (13)) $\bar{d}(\dot{\Gamma}_i^*, \dot{\Gamma}_{i+1}^*)\leqq(\varepsilon/\underline{C})$, $i=0,\cdots,n-1$, where $\mathring{U}_i(p)=1-\varepsilon$ on $\dot{\Gamma}_{i+1}^*$. By applying (15) and (20), we obtain $\bar{d}(\dot{\Gamma}_i^*, \dot{\Gamma}_{i+1}^*)\leqq(\bar{C}(\alpha)/\underline{C}\cdot N(\alpha))\cdot\delta$, $(i+1)\cdot\delta\leqq\alpha$.

Proof of (18). Given $\alpha\in[0, A_0)$, there is a bound $R(\alpha)<\infty$ such that if $i\cdot\delta\leqq\alpha$ and $\delta>0$ is sufficiently small, then $\sup\{|p|: p\in\dot{\Gamma}_i\}<R(\alpha)$. In fact using (13) and (17), we obtain

$$\bar{d}(\partial Q, \dot{\Gamma}_i)\leqq\bar{d}(\dot{\Gamma}_i^*, \dot{\Gamma}_i)+\sum\bar{d}(\dot{\Gamma}_j^*, \dot{\Gamma}_{j+1}^*)\leqq\bar{d}+(\alpha\cdot\bar{C}(\alpha)/\underline{C}\cdot N(\alpha)),$$

where the sum is over $j=0,\cdots,i-1$. For a fixed i (satisfying $(i+1)\cdot\delta\leqq\alpha$), we have $|\nabla\mathring{U}_i(p)|=\mathring{c}$ on $\mathring{\Gamma}_i=\dot{\Gamma}_{i+1}$. If $\bar{\Omega}=\mu\cdot\dot{\Omega}_i$ and $\tilde{\Omega}=\eta\cdot\dot{\Omega}_i$, where $1<\mu<\eta$ are extremal constants (μ maximum, η minimum) such that $\tilde{S}\subset\mathring{S}_i\subset\bar{S}$ (and where $\mu\cdot\Omega:=\{\mu\cdot p: p\in\Omega\}$), then Lemma 4 implies that $\mathring{c}=|\nabla\mathring{U}_i(p)|\leqq|\nabla\tilde{U}(p)|=(c_i/\eta)$, where $p\in\tilde{\Gamma}\cap\dot{\Gamma}_{i+1}$. Similarly, using $p\in\bar{\Gamma}\cap\dot{\Gamma}_{i+1}$, one can show using (15) and the maximum principle that $\mathring{c}=|\nabla\mathring{U}_i(p)|\geqq|\nabla\bar{U}(p)|\cdot\inf\{\dot{U}_i(q): q\in\bar{\Gamma}^*\}\geqq(1-R(\alpha)\cdot\bar{C}(\alpha)\cdot(\mu-1))\cdot(c_i/\mu)$. By combining inequalities, we obtain

$$\eta\leqq\mu\cdot(1-R(\alpha)\cdot\bar{C}(\alpha)\cdot(\mu-1))^{-1}. \tag{21}$$

One can show using the convexity of $\dot{S}_i^*\cup\dot{\Omega}_i$ that $\bar{d}(\dot{\Gamma}_i, \tilde{\Gamma})\leqq(\eta-1)\cdot R(\alpha)$ and that $\underline{d}(\dot{\Gamma}_i, \bar{\Gamma})\geqq(\mu-1)\cdot\underline{r}$ (where $\underline{r}=\underline{d}(0, \partial Q)$). Therefore, if the definition of $N_1(\alpha)>0$ (relative to $S^*\cup\Omega$) is analogous to the definition of $N(\alpha)$ in the proof of (15), then $\mu-1\leqq(\underline{d}(\dot{\Gamma}_i, \bar{\Gamma})/\underline{r})\leqq(\underline{d}(\dot{\Gamma}_i, \dot{\Gamma}_{i+1})/\underline{r})\leqq(\delta/\underline{r}\cdot N_1(\alpha))$. Therefore, if $\delta>0$ is sufficiently small, then (21) simplifies to $\eta\leqq\mu\cdot(1+2R(\alpha)\cdot\bar{C}(\alpha)\cdot(\mu-1))$, and we obtain: $\bar{d}(\dot{\Gamma}_i, \dot{\Gamma}_{i+1})\leqq\bar{d}(\dot{\Gamma}_i, \tilde{\Gamma})\leqq(\eta-1)\cdot R(\alpha)\leqq(\mu-1)\cdot(1+2R(\alpha)\cdot\bar{C}(\alpha))\cdot R(\alpha)\leqq(1+2R(\alpha)\cdot\bar{C}(\alpha))\cdot(R(\alpha)/\underline{r}\cdot N_1(\alpha))\cdot\delta$ for $(i+1)\cdot\delta\leqq\alpha$.

Proof of (19). For i fixed ($i\cdot\delta\leqq\alpha$), let the periodic analytic function G: $(0, 1)\times R\to\dot{\Omega}_i$ be defined as in the proof of Lemma 5(a). Then $\log(c_i\cdot|G'(z)|)$ has an antisymmetric harmonic continuation to $(-1, 1)\times R$. Since $|\nabla\dot{U}_i(G(z))|=1/|G'(z)|$ on $(0, 1)\times R$, (12) and (15) imply (for $\delta>0$ sufficiently small) that $|\log(c_i\cdot|G'(z)|)|\leqq\log(\bar{C}(\alpha)/\underline{C})$ on $(-1, 1)\times R$. Thus $-\log((\bar{C}(\alpha))^2/\underline{C})\leqq\log(|G'(z)|)\leqq\log(\bar{C}(\alpha)/\underline{C}^2)$ on $(-1, 1)\times R$. Uniform bounds for all derivatives of $\log(|G'(z)|)$ on $(-\frac{1}{2}, \frac{1}{2})\times R$ follow from derivatives of the Poisson integral formula for harmonic solutions of the Dirichlet problem in the circle. (19) follows from this.

5. The proof of Theorem 2, Case 1. The main ideas of the proof are sketched in Figs. 3, 4, and 5. Since K_c is a continuous, monotone increasing function of c on $(0, \infty)$ (whereas $|\Omega_c|$ is continuous and monotone decreasing in c), it is sufficient to consider the case where $|\Omega| = A = |\Omega_c|$, i.e. $|S^*\backslash Q| + |S\backslash S_c| = |S_c\backslash S|$. For any $c' \in (0, c]$, we have $|S_c\backslash S| = |S_{c'}\backslash S| + |(S_c\backslash S_{c'})\backslash S|$. Therefore, there is a constant $\hat{c} \in (0, c]$ such that $|S^*\backslash Q| = |S_{\hat{c}}\backslash S|$ and $|S\backslash S_c| = |(S_c\backslash S_{\hat{c}})\backslash S|$. (See Fig. 3.) If $\hat{\Omega} = R^2\backslash(Q \cup S \cup S_{\hat{c}})$, then $|\hat{\Omega}| = |\Omega_c| + |(S_c\backslash S_{\hat{c}})\backslash S| - |S\backslash S_c| = A$. Therefore, if $\hat{\Omega} \neq \Omega_c$, then $\hat{K} > K_c$ by [1, Thm. 12].

To prove (1), it would now suffice to show that $K > \hat{K}$ whenever $\Omega \neq \hat{\Omega}$. We will first prove, under the additional assumption that $\hat{\Gamma}^*$ $(= \partial Q)$ has bounded curvature, that $K \geqq \hat{K}$. Indeed, it suffices to show for each sufficiently small $\varepsilon > 0$ that

$$\ddot{K}_{2\varepsilon} \geqq \hat{K} \tag{22}$$

(since $\ddot{K}_{2\varepsilon} \to K$ as $\varepsilon \to 0+$), where $\ddot{\Omega}_\varepsilon$ is defined such that $\ddot{S}^*_\varepsilon$ and $\ddot{S}_\varepsilon$ are the closed ε-neighborhoods of S^* and S.

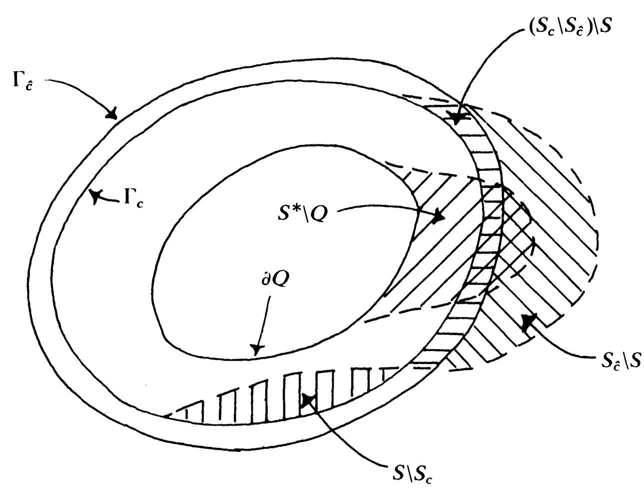

FIG. 3. *$\hat{c} \in (0, c]$ is chosen such that $|(S_c\backslash S_{\hat{c}})\backslash S| = |S\backslash S_c|$ and $|S_{\hat{c}}\backslash S| = |S^*\backslash Q|$.*

To prove (22) (for a fixed, sufficiently small $\varepsilon > 0$), choose $\alpha \in [0, A_0)$ (where $A_0 = |S^*\backslash Q| = |S_{\hat{c}}\backslash S|$) such that the area of the portion of Ω on either side of any straight line intersecting Ω_ε is at least $(A_0 - \alpha)$. Let the regions $\dot{\Omega}_i$, $i = 0, \cdots, n$, be defined (relative to the present regions $\Omega_{\hat{c}}$ and Ω) as in Lemma 6, where $\delta > 0$ and $n \in N$ are chosen such that (a) $n \cdot \delta = A_0$, (b) $\delta \leqq (A_0 - \alpha)/4$, (c) $\delta \leqq \delta(\beta)$, and (d) $M(\beta) \cdot (1 + \bar{C}(\beta)) \cdot \delta \leqq \varepsilon$, and where $\beta = (\alpha + A_0)/2$. ($\bar{C}(\beta)$ and $M(\beta)$ were defined in Lemma 7.) For $\delta > 0$ fixed, we define a sequence of regions $\hat{\Omega}_i$, $i = 0, \cdots, n$, such that $\hat{S}_i = \dot{S}_i \cup \bar{S}$ and $\hat{S}^*_i = \bar{S}^* \cap \dot{S}^*_i$. Here $\bar{S}^*$ and $\bar{S}$ are closed sets chosen such that (i) $S_\varepsilon \subset \bar{S} \subset S_{2\varepsilon}$ and $S^*_\varepsilon \subset \bar{S}^* \subset S^*_{2\varepsilon}$, (ii) the sets $\hat{S}^*_i$, $i = 0, \cdots, n$, contain no isolated points, and (iii) the boundaries $\bar{\Gamma}^*$ and $\bar{\Gamma}$ of $\bar{S}^*$ and $\bar{S}$ are smooth simple closed curves which intersect each of the curves $\dot{\Gamma}^*_i$ and $\dot{\Gamma}_i$, $i = 0, \cdots, n$, at at most a finite number of points.

Although in general $\hat{\Omega}_i$, $i = 0, \cdots, n$, is not doubly connected, the potential $\hat{U}_i(p)$ continues to exist under assumptions (i), (ii), and (iii), and one can define the capacitance of $\hat{\Omega}_i$ by $\hat{K}_i = \int_\gamma D_n \hat{U}_i(p) \cdot |dp|$, where $\gamma \subset \hat{\Omega}_i$ is a finite disjoint union of positively-oriented, smooth simple closed curves such that the collective winding number is 1 about each point in $\hat{S}^*_i$ and 0 about each point in $\hat{S}_i$. Here, the normal derivative $D_n \hat{U}_i(p)$ at $p \in \gamma$ is directed into the region bounded by $\hat{\Gamma}^*_i$ and γ.

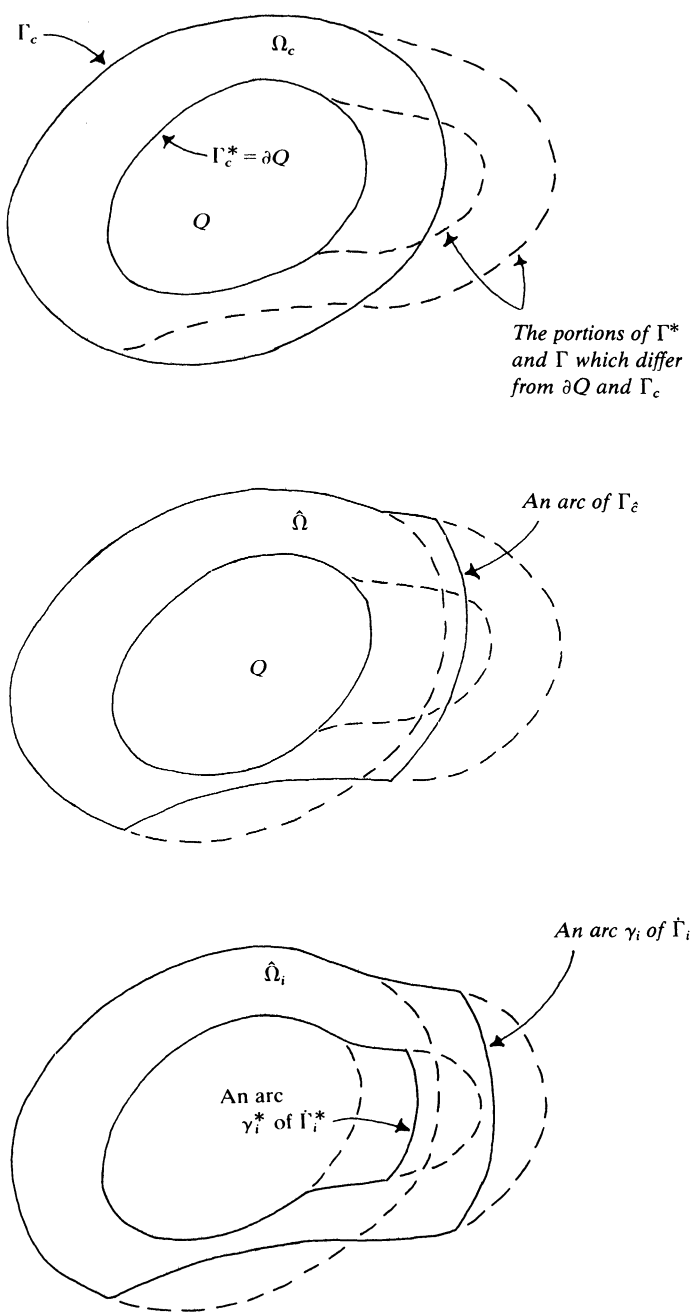

FIG. 4. *Although the proof of Theorem* 2, *Case* 1 *involves further details not shown here, the above diagrams show essentially the stages by which* Ω_c *can be deformed into* Ω *such that the area remains fixed, but the capacitance increases at every step. In the last diagram, the arc segments of* $\dot{\Gamma}_i^*$ *and* $\dot{\Gamma}_i$ *move to the right as i increases until the region* $\hat{\Omega}_i$ *coincides with* Ω.

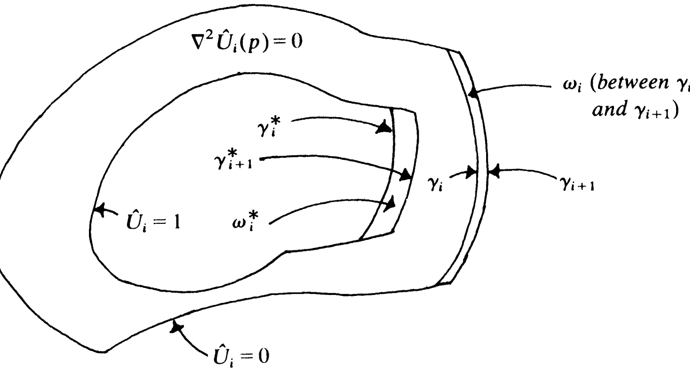

FIG. 5. *Heuristic demonstration that* $\hat{K}_{i+1} \gtrsim \hat{K}_i$: *Since* $|\nabla \dot{U}_i(p)| = c_i$ *on* $\dot{\Gamma}_i$, *Lemma* 5(c) *implies* $|\nabla \dot{U}_i(p)| > c_i$ *on* $\dot{\Gamma}_i^*$. *Then Lemma* 4 *implies* $|\nabla \hat{U}_i(p)| \leqq c_i$ *on* γ_i *and* $|\nabla \hat{U}_i(p)| > c_i$ *on* γ_i^*. *Since* $|\omega_i^*| = |\omega_i|$, *one concludes using the Poincaré variation formula* (3) *that* $\hat{K}_{i+1} - \hat{K}_i \simeq \int_{\gamma_i^*} |\nabla \hat{U}_i(p)|^2 \delta n(p) \cdot |dp| - \int_{\gamma_i} |\nabla \hat{U}_i(p)|^2 \delta n(p) \cdot |dp| \gtrsim c_i^2 \cdot (|\omega_i^*| - |\omega_i|) = 0$.

Now $\hat{S}_0^* = \bar{S}^* \cap Q = Q$ and $\hat{S}_0 = S_{\bar{\varepsilon}} \cup S = \hat{S}$. It follows that $\hat{\Omega}_0 \subset \hat{\Omega}$ and $\hat{K}_0 \geqq \hat{K}$. Further, the above choice of α implies (due to the convexity of $\dot{S}_i^* \cup \dot{\Omega}_i$) that $\dot{S}_i \subset \ddot{S}_\varepsilon$ for $\alpha \leqq i \cdot \delta \leqq A_0$. Therefore $\hat{S}_i = \dot{S}_i \cup \bar{S} \subset \ddot{S}_{2\varepsilon}$ for $\alpha \leqq i \cdot \delta \leqq A_0$. Since $\hat{S}_i^* \subset \ddot{S}_{2\varepsilon}^*$ for all i, we obtain $\ddot{\Omega}_{2\varepsilon} \subset \hat{\Omega}_i$ and $\ddot{K}_{2\varepsilon} \geqq \hat{K}_i$ for $\alpha \leqq j \cdot \delta \leqq A_0$. Combining inequalities yields

$$\ddot{K}_{2\varepsilon} - \hat{K} \geqq \hat{K}_i - \hat{K}_0, \qquad \alpha \leqq i \cdot \delta \leqq A_0. \tag{23}$$

It will be shown that a constant $B(\beta)$ exists such that

$$\Delta_i := \hat{K}_{i+1} - \hat{K}_i \geqq B(\beta) \cdot \delta^2 \tag{24}$$

for $(i+1) \cdot \delta \leqq \beta$, and therefore

$$\hat{K}_i - \hat{K}_0 \geqq \beta \cdot B(\beta) \cdot \delta \tag{25}$$

for $\alpha \leqq i \cdot \delta \leqq \beta$. (An i such that $\alpha \leqq i \cdot \delta \leqq \beta$ exists is due to condition (b) on δ.) (22) follows from (23) and (25), since $\delta > 0$ can be chosen arbitrarily small.

We now prove (24). We have $\Delta_i = I(\Omega_{1,i};\ V_{1,i})$, where $V_{1,i}(p)$ is the harmonic function in $\Omega_{1,i} := \hat{\Omega}_i \cap \hat{\Omega}_{i+1}$ satisfying the boundary conditions $V_{1,i} = \hat{U}_{i+1}$ on $\Gamma_{1,i} = \hat{\Gamma}_i$, $V_{1,i} = 1 - \hat{U}_i$ on $\Gamma_{1,i}^* = \hat{\Gamma}_{i+1}^*$. Here $I(\Omega_{1,i};\ V_{1,i})$ refers to the electric field flux from $S_{1,i}^*$ to $S_{1,i}$ through $\Omega_{1,i}$ due to the potential $V_{1,i}(p)$, i.e. $I(\Omega_{1,i};\ V_{1,i}) = \int_\gamma D_n V_{1,i}(p) \cdot |dp|$, where γ and the sense of $D_n V_{1,i}(p)$, $p \in \gamma$, are fixed as in the definition of $\hat{K}_i$ given earlier. Define $\gamma_i = \dot{\Gamma}_i \backslash \bar{S}$ and $\gamma_i^* = \dot{\Gamma}_i^* \cap \bar{S}^*$. γ_i and γ_i^* are finite unions of arc segments of $\dot{\Gamma}_i$ and $\dot{\Gamma}_i^*$ for each i, due to property (iii) of $\bar{S}$ and $\bar{S}^*$. For $i = 0, \cdots, n-1$ and $j = 2, 3m, 4$, let the harmonic functions $V_{j,i}(p)$ be defined respectively on the regions $\Omega_{2,i} = \Omega_{1,i}$, $\Omega_{3,i} = R^2 \backslash (\hat{S}_{i+1}^* \cup \dot{S}_i)$, and $\Omega_{4,i} = \dot{\Omega}_i \cap \dot{\Omega}_{i+1}$ by the following boundary conditions. $V_{j,i} = \dot{U}_{i+1}$ on γ_i, $V_{j,i} = 1 - \dot{U}_i$ on γ_{i+1}^*, and $V_{j,i} = 0$ on the remaining boundary on $\Omega_{j,i}$. We will show that

$$\Delta_i = I(\Omega_{1,i};\ V_{1,i}) \geqq I(\Omega_{4,i};\ V_{4,i}), \qquad i = 0, \cdots, n-1. \tag{26}$$

Our argument is based on the fact that $I(\Omega;\ V) \geqq 0$ for any $V(p)$ (harmonic in Ω, continuous in Closure (Ω)) satisfying $V(p) \geqq V(q)$ for all $p \in \Gamma^*$, $q \in \Gamma$. It follows from the maximum principle that $V_{1,i}(p) = b1 - \hat{U}_i(p) \geqq 1 - \bar{U}_i(p) \geqq 1 - \dot{U}_i(p) = V_{2,i}(p)$ on γ_{i+1}^* (where $\bar{\Omega}_i = R^2 \backslash (\hat{S}_i^* \cup \dot{S}_i)$), whereas $V_{1,i}(p) \leqq V_{2,i}(p)$ on γ_i by a similar argument. Therefore $I(\Omega_{1,i};\ V_{1,i}) \geqq I(\Omega_{2,i};\ V_{2,i})$. Furthermore, $V_{2,i} = V_{3,i}$ on $\hat{\Gamma}_{i+1}^*$ and $V_{2,i} \leqq V_{3,i}$

on $\hat{\Gamma}_i$, implying that $I(\Omega_{2,i}; V_{2,i}) \geqq I(\Omega_{2,i}; V_{3,i}) = I(\Omega_{3,i}; V_{3,i})$. Finally, $V_{3,i} \geqq V_{4,i}$ on $\dot{\Gamma}^*_{i+1}$ and $V_{3,i} = V_{4,i}$ on $\dot{\Gamma}_i$, implying that $I(\Omega_{3,i}; V_{3,i}) \geqq I(\Omega_{3,i}; V_{4,i}) = I(\Omega_{4,i}; V_{4,i})$.

By applying Green's second identity to the functions $\tilde{U}_i(p)$ and $W_i(p) := V_{4,i}(p)$ in $\tilde{\Omega}_i(0, \lambda)$, $0 < \lambda < 1$ (where $\tilde{\Omega}_i = \Omega_{4,i} = \dot{\Omega}_i \cap \dot{\Omega}_{i+1}$), we obtain

$$I(\tilde{\Omega}_i; W_i) = \int_{\gamma_i^*(\lambda)} D_n W_i(p) \cdot |dp|$$

$$= (1/\lambda) \cdot \int_{\gamma_i^*(\lambda)} \tilde{U}_i(p) \cdot D_n W_i(p) \cdot |dp| \tag{27}$$

$$= (1/\lambda) \cdot \left(\int_{\gamma_i^*(\lambda)} |\nabla \tilde{U}_i(p)| \cdot W_i(p) \cdot |dp| - \int_{\tilde{\Gamma}_i} |\nabla \tilde{U}_i(p)| \cdot W_i(p) \cdot |dp| \right),$$

$i = 0, \cdots, n$, where $\gamma_i^*(\lambda) = \{p \in \tilde{\Omega}_i;\ \tilde{U}_i(p) = \lambda\}$. Lemma 5(c) shows that $|\nabla \tilde{U}_i(p)| \geqq c_i$ in $\tilde{\Omega}_i$, whereas for $(i+1) \cdot \delta \leqq \beta$, (15) and (17) imply that $(1 - \bar{C}(\beta) \cdot M(\beta) \cdot \delta) \cdot |\nabla \tilde{U}_i(p)| \leqq |\nabla \tilde{U}_i(p)| \cdot (\inf \{\dot{U}_i(q) \colon q \in \dot{\Gamma}^*_{i+1}\}) \leqq |\nabla \dot{U}_i(p)| = c_i$ for all $p \in \tilde{\Gamma}_i = \dot{\Gamma}_i$. This reduces to $|\nabla \tilde{U}_i(p)| \leqq (1 + 2\bar{C}(\beta) \cdot M(\beta) \cdot \delta) \cdot c_i$ (when $\bar{C}(\beta) \cdot M(\beta) \cdot \delta < \varepsilon < \frac{1}{2}$), so that for $(i+1) \cdot \delta \leqq \beta$, (27) implies

$$I(\tilde{\Omega}_i; W_i) \geqq (c_i/\lambda) \cdot \left(\int_{\gamma_i^*(\lambda)} W_i(p) \cdot |dp| - (1 + 2\bar{C}(\beta) \cdot M(\beta) \cdot \delta) \cdot \int_{\tilde{\Gamma}_i} W_i(p) \cdot |dp| \right). \tag{28}$$

By combining (26) and (28), taking the limit as $\lambda \to 1-$, and substituting the boundary conditions for $W_i(p) = V_{4,i}(p)$, we obtain

$$\Delta_i \geqq c_i \cdot \left(\int_{\gamma^*_{i+1}} (1 - \dot{U}_i(p)) \cdot |dp| - (1 + 2\bar{C}(\beta) \cdot M(\beta) \cdot \delta) \cdot \int_{\gamma_i} \dot{U}_{i+1}(p) \cdot |dp| \right). \tag{29}$$

We will show for $(i+1) \cdot \delta \leqq \beta$ that

$$J_i^* := \int_{\gamma^*_{i+1}} (1 - \dot{U}_i(p)) \cdot |dp| \geqq c_i \cdot \delta \tag{30}$$

and

$$J_i := \int_{\gamma_i} \dot{U}_{i+1}(p) \cdot |dp| \leqq c'_{i+1} \cdot \delta, \tag{31}$$

where $c'_{i+1} = \sup \{|\nabla \dot{U}_{i+1}(p)| \colon p \in \dot{\Omega}_{i+1} \cap \dot{S}_i\}$. For each i, we have $J_i^* \geqq c_i \cdot \int_{\gamma^*_{i+1}} \underline{d}(p, \dot{\Gamma}^*_i) \cdot |dp|$, since $1 - \dot{U}_i(p) = \int_l |\nabla \dot{U}_i(q)| \cdot |dq| \geqq c_i \cdot L(l) \geqq c_i \cdot \underline{d}(p, \dot{\Gamma}^*_i)$ for each $p \in \gamma^*_{i+1}$. (Here, l is the curve of steepest ascent of $\dot{U}_i(q)$ from p to $\dot{\Gamma}^*_i$.) If σ_i^* is the set swept out by the lines of shortest distance from points $p \in \gamma^*_{i+1}$ to $\dot{\Gamma}^*_i$, then $\int_{\gamma^*_{i+1}} \underline{d}(p, \dot{\Gamma}^*_i) \cdot |dp| \geqq |\sigma_i^*|$, due to the convexity of $\dot{S}^*_i$. One concludes from $\bar{d}(\dot{\Gamma}^*_i, \dot{\Gamma}^*_{i+1}) \leqq M(\beta) \cdot \delta < \varepsilon$ (due to (17) and the condition (d) on δ) and $\ddot{S}^*_\varepsilon \subset \bar{S}^*$ that $(\dot{S}^*_{i+1} \backslash \dot{S}^*_i) \cap S^* \subset \sigma_i^*$, or $|\sigma_i^*| \geqq |(\dot{S}^*_{i+1} \backslash \dot{S}^*_i) \cap S^*| = \delta$. Therefore $J_i^* \geqq c_i \cdot \delta$. For the proof of (31), we define $l(p)$, for any $p \in \gamma_i$, to be the intersection with Closure $(\dot{S}_i \backslash \dot{S}_{i+1})$ of the perpendicular to $\dot{\Gamma}_i$ through p. Also, let $\bar{p} = l(p) \cap \dot{\Gamma}_{i+1}$, $p \in \gamma_i$, and $\sigma_i = \bigcup_{p \in \gamma_i} l(p)$. Now $J_i \leqq c'_{i+1} \cdot \int_{\gamma_i} |p - \bar{p}| \cdot |dp|$, since $\dot{U}_{i+1}(p) \leqq \int_{l(p)} |\nabla \dot{U}_{i+1}(q)| \cdot |dq| \leqq c'_{i+1} \cdot |p - \bar{p}|$ for any $p \in \gamma_i$. Moreover, $\int_{\gamma_i} |p - \bar{p}| \cdot |dp| \leqq |\sigma_i|$, due to the convexity of $\dot{S}^*_i \cup \dot{\Omega}_i$. One concludes from $\bar{d}(\dot{\Gamma}_i, \dot{\Gamma}_{i+1}) \leqq M(\beta) \cdot \delta < \varepsilon$ and $\bar{S} \supset \ddot{S}_\varepsilon$ that $\sigma_i \subset (\dot{S}_i \backslash \dot{S}_{i+1}) \backslash S$, and therefore that $|\sigma_i| \leqq |(\dot{S}_i \backslash \dot{S}_{i+1}) \backslash S| = \delta$. Therefore $J_i \leqq c'_{i+1} \cdot \delta$.

By combining (29), (30), and (31), we obtain (for $(i+1) \cdot \delta \leqq \beta$):

$$\Delta_i \geqq c_i \cdot (c_i - c'_{i+1}) \cdot \delta - 2 c_i \cdot c'_{i+1} \cdot \bar{C}(\beta) \cdot M(\beta) \cdot \delta^2. \tag{32}$$

It follows from (15), (16), (18), and (19) that $0 \leqq c_i \leqq \bar{C}(\beta)$, $0 \leqq c'_{i+1} \leqq \bar{C}(\beta)$, and $c_i - c'_{i+1} = (c_i - c_{i+1}) + (c_{i+1} - c'_{i+1}) \geqq -M(\beta) \cdot (1 + M(\beta)) \cdot \delta$ for $(i+1) \cdot \delta \leqq \beta$. Now (24) follows by substituting these inequalities into (32); in fact we obtain $B(\beta) = -\bar{C}(\beta) \cdot M(\beta) \cdot (1 + M(\beta) + 2\bar{C}^2(\beta))$. This completes the proof that $K \geqq \hat{K}$ when ∂Q has bounded curvature.

To prove that $K \geqq \hat{K}$ in the general case (where Q is convex and ∂Q is a simple closed curve), let $\{Q_n\}$ be an increasing sequence of convex sets such that $\bigcup_{n \in N} Q_n = Q$ and ∂Q_n has bounded curvature for each n. For each n, define Ω_n such that $S_n^* = Q_n$, $|\nabla U_n(p)|$ is constant on Γ_n, and $|S_n \backslash S| = |S^* \backslash Q_n|$. We have $K \geqq \hat{K}_n$ for each n, where $\hat{\Omega}_n = R^2 \backslash (Q_n \cup S \cup S_n)$. Thus $K \geqq \lim_{n \to \infty} \hat{K}_n = \hat{K}$.

Since $\hat{K} > K_c$ when $\hat{\Omega} \neq \Omega_c$, it follows from $K \geqq \hat{K}$ that $K \geqq K_c$, where equality can occur at most in the case where $S \subset S_c$ ($\Rightarrow \hat{c} = c$ and $\hat{\Omega} = \Omega_c$). In order to complete the proof of (1), we will now show that $K > K_c$ whenever $\Omega \neq \Omega_c$ and $S \subset S_c$. By Lemma 5(c), there is an $\varepsilon > 0$ such that $|\nabla U_c(p)| > c + 2\varepsilon$ in $\Omega_c(\frac{1}{4}, 1)$. Therefore, one can choose $b < c$ so close to c that $|\nabla U_b(p)| > c + \varepsilon$ in $\Omega_b(\frac{1}{2}, 1)$ and $\frac{1}{2} < \mu < 1$, where μ is defined by $|S^* \cap \Omega_b(\mu, 1)| = |(S_c \backslash S_b) \backslash S|$. For $0 \leqq \lambda \leqq 1$, define $\bar{\Omega}_{b,\lambda} = R^2 \backslash (S^*_{b,\lambda} \cup S_b \cup S)$ and $\tilde{\Omega}_{b,\lambda} = R^2 \backslash ((S^* \cap S^*_{b,\lambda}) \cup S_b \cup S)$, where $\Omega_{b,\lambda} = \Omega_b(0, \lambda)$. [1, Thm. 8] implies that

$$\tilde{K}_{b,1} \geqq K_c - c^2 \cdot |(S_c \backslash S_b) \backslash S|. \tag{33}$$

Furthermore, for any $\mu \leqq \lambda < 1$ and $p \in \Gamma^*_{b,\lambda} = \bar{\Gamma}^*_{b,\lambda}$, we have using Lemma 4 that $|\nabla \bar{U}_{b,\lambda}(p)| \geqq |\nabla U_{b,\lambda}(p)| = (|\nabla U_b(p)|/\lambda) \geqq c + \varepsilon$. Therefore, [1, Thm. 1] can be applied to show that

$$\tilde{K}_{b,\mu} \geqq \tilde{K}_{b,1} + (c + \varepsilon)^2 \cdot |S^* \cap \Omega_b(\mu, 1)|. \tag{34}$$

It follows from (33) and (34) that

$$\tilde{K}_{b,\mu} - K_c \geqq 2c \cdot \varepsilon \cdot |(S_c \backslash S_b) \backslash S| > 0.$$

Therefore $K > K_c$, since for $S \subset S_c$ the proof that $K \geqq \tilde{K}_{b,\mu}$ is essentially identical to the proof already given that $K \geqq \hat{K}$.

Remark 5. On the basis of Theorem 2, the following conjecture appears reasonable. Let a continuously-differentiable function $0 < r(p) \leqq 1$ represent the electrical resistivity in $R^2 \backslash Q$, where Q is convex. Define the current leakage from Q in the presence of the function $r(p)$ by $I = \int_\gamma (D_n U(p)/r(p)) \cdot |dp|$, where the potential $U(p)$ in $R^2 \backslash Q$ solves the boundary value problem: $\nabla \cdot (\nabla U(p)/r(p)) = 0$ in $R^2 \backslash Q$, $U(p) = 1$ on ∂Q, and $\lim_{|p| \to \infty} U(p) = 0$. (Here, γ is a smooth simple closed curve encircling Q and the normal derivative is inwardly directed.) Then $I \geqq K_c$, where $c > 0$ is chosen such that $|\Omega_c| = \int_{R^2 \backslash Q} r(p)\, dx\, dy$.

Note added in proof. The procedure used in this paper can be extended to prove a generalization of Theorem 2 to be discussed in detail in [8]. In order to state this result, we define $\|\Omega\| = \iint_\Omega a^2(p)\, dx\, dy$, where $a(p) > 0$ is a continuous function on R^2.

THEOREM. *Assume that $a(p)$ is subharmonic in $R^2 \backslash Q$, that ∂Q has bounded curvature, and that there is a point $p_0 \in Q$ such that Q is starlike relative to p_0 and $\lambda \cdot a(p_0 + \lambda \cdot (p - p_0))$ is weakly increasing in $\lambda \in [1, \infty)$ for each $p \in \partial Q$. Then*:

(a) *If Ω_c is defined such that $S_c^* = Q$ and $|\nabla U_c(p)| = c \cdot a(p)$ on Γ_c for each $c \in R_+ = (0, \infty)$, then there exists a constant $0 \leqq C_0 \leqq \infty$ such that $|\nabla U_c(p)| \geqq c \cdot a(p)$ on ∂Q only for those $c \in R_+$ satisfying $c \leqq C_0$.*

(b) *If $c \in [0, C_0] \cap R_+$, then*

$$K \geqq K_c \tag{35}$$

for any Ω satisfying $S^* \supset Q$ and $\|\Omega\| \leqq \|\Omega_c\|$.

(c) *If $C_0 < c < \infty$, then* (35) *does not hold for all admissible Ω.*

REFERENCES

[1] A. ACKER, *Heat flow inequalities with applications to heat flow optimization problems*, this Journal, 8 (1977), pp. 604–618.

[2] ———, *Isoperimetric inequalities involving heat flow under linear radiation conditions*, Proc. Amer. Math. Soc., 64 (1977), pp. 265–272.

[3] ———, *An isoperimetric inequality involving conformal mapping*, Ibid., 65 (1977), pp. 230–234.

[4] ———, *A free boundary optimization problem involving weighted areas*, Z. Agnew. Math. Phys., 29 (1978), pp. 395–408.

[5] ———, *Another free boundary optimization problem involving weighted areas*, Ibid., 29 (1978), pp. 409–413.

[6] ———, *Some free boundary optimization problems and their solutions*, Numerische Behandlung von Differentialgleichungen mit besonderer Berücksichtigung freier Randwertaufgaben, J. Albrecht, L. Collatz, G. Hämmerlin, eds., Birkhaüser Verlag Basel, 1978, pp. 9–22.

[7] ———, *Isoperimetric inequalities involving heat flow—the case of a one-sided linear radiation condition*, Z. Angew. Math. Phys., to appear.

[8] ———, *A free boundary optimization problem—II*, submitted for publication.

[9] A. BEURLING, *On free boundary problems for the Laplace equation*, Seminars on Analytic Functions, Institute for Advanced Study, Princeton, NJ, 1 (1957), pp. 248–263.

[10] T. CARLEMAN, *Über ein Minimalproblem der mathematischen Physik*, Math. Zeitschrift, 1 (1918), pp. 208–212.

[11] M. A. LAVRENTEV, *Variational Methods*, Noordhoff, Groningen, The Netherlands, 1963.

[12] L. E. PAYNE, *Isoperimetric inequalities and their applications*, SIAM Rev., 9 (1967), pp. 453–488.

[13] G. PÓLYA, *Torsional rigidity, principal frequency, electrostatic capacity and symmetrization*, Quart. Appl. Math., 6 (1948), pp. 267–277.

[14] G. PÓLYA AND G. SZEGÖ, *Inequalities for the capacity of a condenser*, Amer. J. Math., 67 (1945), pp. 1–32.

[15] ———, *Isoperimetric Inequalities in Mathematical Physics*, Princeton University Press, Princeton, NJ, 1951.

[16] M. M. SCHIFFER, *Partial differential equations of the elliptic type*, Lecture series of the symposium on partial differential equations held at the University of California at Berkeley (June 20–July 1, 1955), pp. 97–149.

[17] G. SZEGÖ, *Über einige Extremalaufgaben der Potentialtheorie*, Math. Zeitschrift, 31 (1930), pp. 583–593.

[18] D. E. TEPPER, *Free boundary problem*, this Journal, 5 (1974), pp. 841–846.

[19] ———, *Free boundary problem—the starlike case*, this Journal, 6 (1975), pp. 503–505.